

# **Re-finding Tweets**

Analyse der Personal-Information-Management-Praktik Re-finding  
im Kontext der Social-Media-Plattform Twitter



Inaugural-Dissertation zur Erlangung der Doktorwürde der Fakultät für Sprach-,  
Literatur- und Kulturwissenschaften der Universität Regensburg

Vorgelegt von  
**Florian Meier**  
aus Raining bei Haarbach  
Regensburg 2017

**Abgabe:** April 2017

**Tag der Disputation:** 18. Mai 2018

**Erstgutachter:** Prof. Dr. David Elsweiler

**Zweitgutachter:** Prof. Dr. Christian Wolff







# Danksagung

Eine Doktorarbeit zu verfassen und einen Marathon zu laufen weisen viele Parallelen auf. Für beides benötigt man Training, Hingabe und vor allem Ausdauer. Von außen betrachtet mag beides als etwas erscheinen, das man *alleine* bestreitet. Klar ist jedoch auch: Hinter einem erfolgreichen Marathonläufer steht immer auch ein Team, das ihn unterstützt. Ein Team, das je nach den Ambitionen des Läufers kleiner oder größer ausfallen kann. Ein Hobbyläufer mag unter Umständen nur ein kleines Team besitzen, das vielleicht aus einem liebevollen Mann besteht, der dem neuen Hobby und Ziel der Lebenspartnerin verständnisvoll gegenüber steht. Und der kleinen Tochter, die ihre Mutter bei längeren Sonntagsläufen auf dem Rad begleitet. Das Team eines professionellen Läufers mag hingegen deutlich größer ausfallen. Dieser besitzt sicher einen oder sogar mehrere Trainer, medizinische Betreuer, vielleicht einen Ernährungsberater und einen Ausrüster, der ihm maßgeschneiderte Schuhe zur Verfügung stellt. Eine Dissertation trägt nur den Namen eines Autors. Sie wäre aber nicht möglich ohne die Unterstützung des Teams, das hinter dem Verfasser steht. Ich hatte das große Glück, während meiner ganzen Promotionszeit das professionellste und beste Team an meiner Seite zu haben, das man sich nur vorstellen kann. In den folgenden Zeilen möchte ich mein Team vorstellen und ihm meinen Dank aussprechen.

Ganz besonders bedanken möchte ich mich bei meinen Betreuern und Gutachtern Prof. Dr. David Elweiler und Prof. Dr. Christian Wolff. Die Betreuung, die ich durch David erfahren habe, kann nur als absolut wundervoll bezeichnet werden. Er behauptet zwar immer scherzhaft, dass er mehr als Psychologe fungiert habe und es mehr seine Aufgabe war, mich wieder aufzurichten, wenn ich demotiviert und enttäuscht war, als den Verlauf meines Projekts zu steuern. Tatsache ist jedoch: Die Arbeit hätte ohne ihn nie so entstehen können. Ich habe in diesen Jahren der Dissertation so viele neue Dinge gelernt, die ich ohne ihn vielleicht nie in Betracht gezogen hätte. Fast wichtiger als die entstandene Dissertation ist jedoch, dass neben dieser auch noch eine Freundschaft gewachsen ist. Ich bin sehr dankbar und glücklich darüber, David nun als meinen Freund bezeichnen zu können und freue mich auf weitere Zusammenarbeit. Herr Wolff fungierte bereits während meines Studiums als Förderer und Mentor für mich. Nur durch seine Ambitionen konnte die von Sarah Will und mir verfasste Projektarbeit zur Entwicklung des Webdesigns als wissenschaftlicher Artikel im Rahmen der Mensch und Computer 2009 vorgestellt und veröffentlicht werden. Er hat mir die akademische Welt schmackhaft gemacht und besitzt somit großen Anteil an der Umsetzung der Arbeit und meinem Werdegang im Allgemeinen. Bedanken möchte ich mich in diesem Zuge auch bei Dr. Morgan Harvey, der als zusätzlicher Gutachter für meine Arbeit fungiert hat. Schließlich möchte ich nicht vergessen, an Herrn Prof. Dr. Rainer Hammwöhner zu erinnern, der mir damals die Möglichkeit gab, eine Arbeitsstelle als Wissenschaftlicher Mitarbeiter am Lehrstuhl für Informationswissenschaft anzutreten und den ich für seinen wissenschaftlichen Scharfsinn und seinen Umgang mit Menschen stets beachtet habe.

Viele meiner Freunde haben direkt oder indirekt an dieser Arbeit mitgewirkt. Hiermit bedanke ich mich recht herzlich bei: Alex Bazo und Thomas 'Kaiser' Wilhelm, die mir in vielerlei Hinsicht stets

geholffen haben, sei es mit Problem bei der Implementierung der Chrome-Extension, bei der Analyse von qualitativen Daten oder durch das Korrekturlesen der Arbeit. Ich könnte mich nicht daran erinnern, dass ich jemals ein *Nein* von ihnen gehört hätte, wenn ich sie um Hilfe gebeten habe; Max L. Wilson, mit dem ich die Ehre hatte während meines Aufenthalts als wissenschaftlicher Praktikant in Nottingham zusammenzuarbeiten, der die finanziellen Mitteln für die Fragebogenstudie aufgebracht hat und bei der Analyse der qualitativen Daten beteiligt war; Markus Kattenbeck, der stets ein offenes Ohr für mich hatte, um mit mir über statistische Problem zu sprechen und Lösungen zu finden; Isabella Hastreiter, die mir bei der Auswertung der qualitativen Daten geholfen hat; Manuel Burghardt, der ständig gute Tipps für das Verfassen einer Doktorarbeit hatte und mir seine Latex-Vorlage überlassen hat. Schließlich müssen auch die unerschrockenen Korrekturlesenden erwähnt werden. Danke Seppe, Sarah und Bärbel für eure Zeit, die ihr sicherlich auch mit etwas Angenehmeren verbringen hättet können.

Durch meine Arbeit am Lehrstuhl für Informationswissenschaft durfte ich über den Zeitraum der fast sechs Jahre, die ich dort gearbeitet habe, viele nette Kollegen am eigenen und an benachbarten Lehrstühlen kennenlernen. Diese haben jeden Arbeitstag, war er auch noch so anstrengend, stets erträglich gemacht. Mein Dank gilt: Barbara, Bernd, Bienki, Christina, Daniel, Hanna, Ingrid, Jürgen, Manuel, Markus F., Martin, Melanie, Florin, Patricia, Rafael, Ray, Renate, Robert, Tim, Valentin, Victoria ... there are too many to mention.

Ganz besonderer Dank gebührt natürlich meiner Familie, besonders meinen Eltern und meiner Schwester Anna, die mich stets unterstützt haben, die nie einen Schritt in meinem Leben hinterfragt oder kritisiert haben und mich 'einfach haben machen lassen'. Besonders zu erwähnen sind hier mein Onkel Hans und meine Tante Bärbel, die schon während meines Studiums als Korrekturlesende geholfen haben und mir gezeigt haben, wie man eine richtige wissenschaftliche Arbeit verfasst.

Ich möchte nicht zum Ende kommen, ohne meinen langjährigen Freunden aus dem Ö-Stammtisch zu danken. Teil meiner Motivation eine Doktorarbeit zu verfassen war sicherlich auch darin begründet, dass mehr Zeit an der Universität auch mehr Zeit mit euch bedeutete. Danke Chrisi, Wuttke, Pocky, Hosh, Dennis, Tobi und Balki für eine unvergessliche Zeit :).

Tusind tak skal I have allesammen!!!

Kopenhagen im Dezember 2018

# Zusammenfassung

Diese Arbeit untersucht das Informationsverhalten von Social-Media-Anwendern aus der Perspektive des *Personal Information Management* und fokussiert dabei auf Re-finding-Verhalten, also das Wiederfinden von bereits wahrgenommener Information. Als Untersuchungsgegenstand dient die Social-Media-Plattform Twitter. Ziel der Arbeit ist die Beobachtung, Dokumentation, Beschreibung und Interpretation des Nutzerverhaltens beim Wiederfinden von Tweets und die Erarbeitung von Designvorschlägen, um Twitter-Nutzer bei diesem Informationsbedürfnis zu unterstützen. Als Forschungsstrategie dient ein Sequential-Mixed-Methods-Design, welches die sukzessive Erhebung und Auswertung von qualitativen bzw. subjektiven und quantitativen bzw. objektiven Daten in Form von zwei großen Studien — einer Umfrage und einer Logstudie — ermöglicht und es schließlich erlaubt, durch Kombination und Diskussion der Einzelergebnisse ein holistisches Bild von Wiederfindensverhalten auf Twitter zu zeichnen. Die Arbeit zeigt, dass Nutzer sehr häufig das Bedürfnis haben, zu bereits gesehenen Tweets zurückzukehren. Twitter, obwohl es einen Fokus auf Echtzeitinformationen legt, besitzt Archivcharakter, da häufig auch ältere Nachrichten wieder aufgerufen werden und persönliche Tweets einen längeren Lebenszyklus besitzen, als man dies von ihnen erwarten würde. Wiederfindensstrategien — besonders Orienteering-Verhalten — die bereits in anderen Personal-Information-Management-Kontexten wie mit E-Mails oder bei der Nutzung von Dateimanagern identifiziert werden konnten, treten auch beim Wiederfinden von Tweets auf. Wiederfinden kann eine komplexe Aufgabe sein, die Nutzer frustriert zurücklässt. Darüber hinaus haben Nutzer Schwierigkeiten bei der Einschätzung, ob Tweets in Zukunft von Relevanz sein könnten. Angemessen trainierte Algorithmen können Nutzer beim Wiederfinden von Tweets unterstützen.



# Abstract

The thesis investigates social media information behaviour from a personal information management point of view while focussing on re-finding behaviour. The social media platform Twitter serves as an object of investigation. The thesis aims at observing, documenting, describing and interpreting user behaviour when re-finding Tweets. Based on these observations design implications for supporting users exhibiting this behaviour are presented. A sequential *Mixed Methods* research design is employed, which allows for successively collecting and analysing qualitative and quantitative data using two approaches: a large-scale survey and a naturalistic log data analysis. The combination and discussion of results from both studies opens up a holistic view of re-finding behaviour on Twitter. The work shows that Twitter users frequently have the need to re-find previously seen Twitter content. Despite Twitter's emphasis on real-time information the service also exhibits characteristics of an archive as a lot of older content is viewed and the life-cycle of Tweets for individual users is much longer than one might expect. Re-finding strategies — especially orienteering — which were identified as means to re-find by empirical work in other PIM contexts are also adopted in the context of re-finding on Twitter. Re-finding can be hard work leading to frustration. Moreover users have problems with anticipating future information needs with Tweets. This work shows that appropriately trained standard algorithms can support users with this task and thus facilitate re-finding.



# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>1</b>
1.1. Motivation und Zielsetzung . . . . .	2
1.2. Arbeitsdefinitionen . . . . .	3
1.3. Die Social-Media-Plattform Twitter . . . . .	5
1.4. Gliederung der Arbeit . . . . .	11
<b>2. Personal Information Management</b>	<b>13</b>
2.1. Vom Informationsverhalten zur Verwaltung persönlicher Information . . . . .	13
2.1.1. Ein Blick auf Social Media aus der Perspektive der Informationsverhaltensforschung . . . . .	15
2.1.2. Tweets als persönliche Information . . . . .	17
2.1.3. Informationscharakteristika eines Tweets . . . . .	18
2.2. PIM als Forschungsfeld und sein Forschungsstand . . . . .	20
2.2.1. Dimensionen der PIM-Forschung . . . . .	21
2.2.2. PIM-Praktik Aufbewahren . . . . .	24
2.2.3. PIM-Praktiken Organisieren und Verwalten . . . . .	27
2.2.4. PIM-Praktik Wiederfinden . . . . .	30
2.2.4.1. Wiederfinden von E-Mails . . . . .	31
2.2.4.2. Wiederfinden und Wiederbesuch von Webseiten . . . . .	34
2.3. PIM und PIM-Praktiken: Ein Fazit . . . . .	36
<b>3. Social-Media-Forschung: Twitter und Social-Media-PIM im Fokus</b>	<b>39</b>
3.1. Twitter-Forschung im Überblick . . . . .	39
3.2. Social-Media-Informationsverhalten mit Fokus auf PIM-Praktiken . . . . .	45
3.2.1. Social-Media-Inhalte Finden . . . . .	45
3.2.1.1. Gerichtete Suche, Relevanzkriterien und Microblogsuche . . . . .	45
3.2.1.2. Social Media als soziale Frage-Antwort-Systeme . . . . .	48
3.2.1.3. Informationsdiffusion, RT-Vorhersage, Tweet-(Re)-Ranking und Serendipity . . . . .	49
3.2.1.4. Informationserschließung und Hashtagnutzung . . . . .	51
3.2.2. Social-Media-Nachrichten Verwalten . . . . .	53
3.2.2.1. Social Media als Archiv . . . . .	55
3.2.2.2. Social-Media-Inhalte Aufbewahren . . . . .	57
3.2.3. Social-Media-Inhalte Wiederfinden . . . . .	58
3.3. Zusammenfassung: Twitter und Social-Media-PIM . . . . .	59
<b>4. Forschungsdesign</b>	<b>61</b>
4.1. Triangulation und <i>Mixed Methods</i> in der Informationswissenschaft . . . . .	61

4.2. Triangulation und Mixed-Methods-Forschungsdesigns im Detail . . . . .	62
4.3. Das Forschungsdesign im Überblick . . . . .	64
4.3.1. Twitter-Umfrage . . . . .	64
4.3.2. Twitter-Logstudie . . . . .	67
4.3.3. Meta-Inferenz . . . . .	68
<b>5. Studie I: Twitter-Umfrage</b> . . . . .	<b>69</b>
5.1. Leitfragen der Twitter-Umfrage . . . . .	70
5.2. Fragebogendesign und Fragebogengliederung . . . . .	71
5.2.1. Makroplanung des Fragebogens . . . . .	72
5.2.2. Mikroplanung des Fragebogens und Frage-Antwort-Typen . . . . .	72
5.3. Samplingverfahren und webbasierte Rekrutierungs- strategien für Umfragen . . . . .	76
5.3.1. Die Verwendung von Crowdsourcing zur Rekrutierung von Probanden für Fragebogenstudien . . . . .	77
5.3.2. Die Crowdsourcing-Umfrage-Plattform Tellwut . . . . .	81
5.4. Methodendiskussion: Qualitative Auswertungsmethoden im Vergleich . . . . .	82
5.4.1. Die Methode <i>Grounded Theory</i> . . . . .	85
5.4.2. Die Methode qualitative Inhaltsanalyse . . . . .	89
5.4.3. Beschreibung des Vorgehens bei der Datenauswertung . . . . .	90
5.5. Charakterisierung der Studienteilnehmer . . . . .	93
5.6. Die Rolle der Favoriten-Funktion für das PIM-Verhalten auf Twitter . . . . .	96
5.6.1. Der Auswertungsprozess im Detail . . . . .	96
5.6.2. Nutzungshäufigkeit der Favoriten-Funktion . . . . .	98
5.6.3. Motivation und Beweggründe zur Nutzung der Favoriten-Funktion . . . . .	99
5.6.3.1. Kodekategorie [A]: <i>Favouriting as a response/reaction</i> . . . . .	100
5.6.3.2. Kodekategorie [B]: <i>Favouriting for a purpose/as a function</i> . . . . .	103
5.6.3.3. Kodekategorie [C]: <i>No reason but interesting behaviour</i> . . . . .	105
5.6.3.4. Diskussion der Favorisierungsgründe . . . . .	106
5.7. Quantitative Ergebnisse zu Aufbewahrens- und Wiederfindensstrategien . . . . .	109
5.7.1. Signifikanzprüfung mithilfe des nicht-parametrischen Bootstraps . . . . .	110
5.7.2. Aufbewahrenshäufigkeit und -strategien . . . . .	112
5.7.3. Wiederfindenshäufigkeit und -strategien . . . . .	113
5.7.4. Aufbewahrenshäufigkeit und Wiederfindensstrategien . . . . .	115
5.7.5. Schwierigkeit des Wiederfindens und Wiederfindensstrategien . . . . .	116
5.7.6. Bewertung der Twitteroptionen und Wiederfindensstrategien . . . . .	117
5.7.7. Frustrationserfahrung und Wiederfindensstrategien . . . . .	117
5.7.8. Einflussfaktoren auf das Frustrationsniveau bei Wiederfindensaktionen . . . . .	120
5.7.8.1. Logistische Regression als Analysemethode . . . . .	120
5.7.8.2. Die Divide-by-4-Regel . . . . .	121
5.7.8.3. Bestimmung der Anpassungsgüte . . . . .	122
5.7.8.4. Beschreibung der verwendeten Features . . . . .	122
5.7.8.5. Datenanalyse und Modellbildung . . . . .	122
5.7.8.6. Ergebnis: Logistisches Regressionsmodell . . . . .	123
5.7.8.7. Fazit zum Regressionsmodell . . . . .	125
5.8. Analyse von Re-finding-Motiven: Warum Wiederfinden? . . . . .	126
5.8.1. Motivation und Beweggründe für das Wiederfinden von Tweets . . . . .	127
5.8.2. Diskussion der Wiederfindensgründe und Fazit . . . . .	133
5.9. Fazit zur Fragebogenstudie . . . . .	135



<b>6. Studie II: Twitter-Logstudie</b>	<b>137</b>
6.1. Leitfragen der Twitter-Logstudie . . . . .	138
6.2. Logdatenanalyse als Methode . . . . .	139
6.2.1. Logdatenanalyse im informationswissenschaftlichen Forschungskontext . . . . .	140
6.2.2. Charakterisierung von Logdaten . . . . .	143
6.2.3. Methodische Charakterisierung der Logdatenanalyse . . . . .	145
6.2.4. Anwendungsdimensionen und Phasen der Logdatenanalyse . . . . .	147
6.2.5. Überblick zu Studien mit Social-Media-Clickstream-Daten . . . . .	149
6.3. Erhebung der Logdaten für die Twitter-Logstudie . . . . .	151
6.3.1. Chrome Webbrowser-Erweiterung . . . . .	151
6.3.2. <i>Experience Sampling</i> . . . . .	156
6.3.2.1. Die <i>Experience-Sampling</i> -Methode . . . . .	156
6.3.2.2. <i>Experience Sampling</i> in der Twitter-Logerweiterung . . . . .	158
6.3.3. Teilnehmengewinnung und Rekrutierungsstrategien . . . . .	159
6.4. Ergebnisse der Logdatenauswertung . . . . .	160
6.4.1. Vorarbeiten bei der Datenauswertung und -bereinigung . . . . .	160
6.4.1.1. Eventkategorisierung . . . . .	161
6.4.1.2. Einteilung der Daten in Sessions . . . . .	162
6.4.2. Charakterisierung der Studienteilnehmer . . . . .	164
6.4.3. Twitter-Verhalten im Allgemeinen . . . . .	166
6.4.3.1. Twitter-Verhalten in Bezug auf Sessions . . . . .	167
6.4.3.2. Twitter-Verhalten in Bezug auf Eventkategorien . . . . .	168
6.4.3.3. Twitter-Verhalten in Bezug auf Interaktionen mit Tweets . . . . .	171
6.4.3.4. Zusammenfassung und Erkenntnisse für die Re-finding-Analyse . . . . .	173
6.4.4. Analyse des Re-finding-Verhaltens . . . . .	174
6.4.4.1. Klassifikation von Re-finding-Sitzungen . . . . .	175
6.4.4.2. Proxy-Diskussion: Häufigkeit des Auftretens von Re-finding-Verhalten	175
6.4.4.3. Der nutzerbezogene Lebenszyklus eines Tweets . . . . .	177
6.4.4.4. Charakterisierung des Re-finding-Verhaltens auf Sessionebene . . . . .	179
6.4.4.5. Analyse von Wiederfindensstrategien . . . . .	182
6.4.4.6. Analyse von Re-finding-Pfaden . . . . .	184
6.4.4.7. Analyse von Wiederfindensgründen . . . . .	186
6.4.4.8. Analyse von erfolglosem Re-finding-Verhalten . . . . .	186
6.4.5. Vorhersage von Re-finding-Zielen . . . . .	189
6.4.5.1. Experimentelles Setup und <i>Random Forest</i> als Klassifikationsverfahren	190
6.4.5.2. Beschreibung der verwendeten Features . . . . .	192
6.4.5.3. Ergebnisse des Vorhersage-Experiments . . . . .	195
6.4.5.4. Fazit zum Vorhersage-Experiment . . . . .	199
6.5. Fazit zur Twitter-Logstudie . . . . .	200
<b>7. Diskussion: Erkenntnisse aus Literatur und Studien vereint</b>	<b>203</b>
7.1. Fragestellungen zur PIM-Aktivität Aufbewahren . . . . .	203
7.1.1. Aufbewahrensbedürfnis und Strategien des Aufbewahrens . . . . .	204
7.1.2. Twitter als Archiv . . . . .	205
7.2. Fragestellungen zur PIM-Aktivität Wiederfinden . . . . .	206
7.2.1. Relevanz von Tweets: Jetzt! Und in Zukunft? . . . . .	208
7.2.2. Wiederfindensstrategien: <i>Teleporting</i> vs. <i>Orienteering</i> . . . . .	210
7.3. Zusammenfassung: Identifizierte Probleme . . . . .	213

<b>8. Wiederfindensverhalten unterstützen: Lösungsansätze und Designvorschläge</b>	<b>215</b>
8.1. Erinnerungsfunktion für aufbewahrte Tweets . . . . .	216
8.2. Re-finding-Verhalten unterstützen . . . . .	218
8.2.1. <i>Interleaving</i> von Timeline-Tweets mit Wiederfindenszielen . . . . .	218
8.2.2. Verlauf geklickter Tweets . . . . .	219
8.2.3. Unterstützung des Teleporting-Verhaltens: <i>Learning to re-find</i> . . . . .	219
8.2.4. Unterstützung des Orienteering-Verhaltens . . . . .	221
8.3. Fazit zu den Designvorschlägen . . . . .	223
<b>9. Fazit und Ausblick</b>	<b>225</b>
9.1. Hauptbeiträge der Arbeit in Zusammenfassung . . . . .	225
9.2. Kritische Anmerkungen zur Arbeit . . . . .	227
9.2.1. In Bezug auf die verwendeten Methoden . . . . .	227
9.2.2. In Bezug auf den Untersuchungsgegenstand . . . . .	228
9.3. Anknüpfungspunkte für zukünftige Forschung . . . . .	229
9.3.1. In Bezug auf existierende Daten . . . . .	229
9.3.2. In Bezug auf neue Studien . . . . .	230
<b>Literaturverzeichnis</b>	<b>233</b>
<b>Anhang</b>	<b>I</b>
<b>A. Twitter-Umfrage</b>	<b>II</b>
A.1. Fragebogen . . . . .	III
A.2. Signifikanzprüfungen der quantitativen Auswertungen des Fragebogens . . . . .	IX
<b>B. Twitter-Logstudie</b>	<b>X</b>
B.1. Ergebnisse der Shapiro-Wilk-Tests . . . . .	X
B.2. Signifikanzprüfungen beim Vergleich von RF-Sessions mit normalen Sessions . . . . .	XI
B.3. Signifikanzprüfungen beim Vergleich von erfolgreichen und erfolglosen RF-Sessions mit normalen Sessions . . . . .	XIII
B.4. Kodierschema für Tweet-Inhaltskategorien . . . . .	XV

# Abbildungsverzeichnis

1.1. Erster Sketch des Twitter-Vorgänger <i>Stat.us</i> von Jack Dorsey (Dorsey, 2006) . . . . .	5
1.2. Erstmalige Verwendung des Hashtags auf Twitter im August 2007 durch den Nutzer Chris Messina (Screenshot von: <a href="https://twitter.com/chrimessina/status/223115412">https://twitter.com/chrimessina/status/223115412</a> ) . . . . .	7
1.3. Screenshot der Twitter-Suche respektive der Twitter-SERP im April 2017 (Screenshot von: <a href="https://twitter.com/search?src=typd&amp;q=%23corgi">https://twitter.com/search?src=typd&amp;q=%23corgi</a> ) . . . . .	8
2.1. Mentale Fragen, die aus Information persönliche Information werden lassen . . . . .	18
3.1. Anzahl an Veröffentlichungen (pro Jahr) zu bekannten Social-Media-Applikationen in der Literaturdatenbank Scopus . . . . .	40
3.2. Publikationen zum Thema Twitter aufgeschlüsselt nach Veröffentlichungsjahr und Fachgebiet (Export von: <a href="https://www.scopus.com/">https://www.scopus.com/</a>   Stand: November 2016) . . . . .	43
4.1. Visualisierung des Forschungsdesigns auf der Basis des Sequential-Mixed-Design nach Teddlie und Tashakkori (Teddlie & Tashakkori, 2006) . . . . .	65
5.1. Verbalisierte Sieben-Punkt-Antwortskala wie sie bei Fragen zur Häufigkeit von Verhaltensweisen genutzt wird . . . . .	73
5.2. Verbalisierte Fünf-Punkt-Zustimmungsskala wie sie bei Fragen nach konkreten Verhaltensweisen genutzt wird . . . . .	74
5.3. Begrüßungs-Screen der webbasierten Variante der Umfrage . . . . .	75
5.4. Crowdsourcing Landscape Beta v2 nach Dawson und Bynghall (Dawson & Bynghall, 2012) . . . . .	78
5.5. Das Ranking amerikanischer Umfrage-Plattformen auf <i>SurveyPolice.com</i> . Tellwut belegt dort im November 2016 Rang 4. (Screenshot von: <a href="https://www.surveypolice.com/countries/united-states">https://www.surveypolice.com/countries/united-states</a> ) . . . . .	82
5.6. Ausschnitt der Umfrage für die Tellwut-Probanden . . . . .	83
5.7. Charakterisierung von qualitativen Forschungsmethoden nach Tesch (Tesch, 1990) . . . . .	84
5.8. Kodierprozess der <i>Grounded Theory</i> in systematischer Anordnung nach Breuer (Breuer, 2009) . . . . .	88
5.9. Visualisierung des „Werkzeugkastenmodells der Inhaltsanalyse“. Eigene Darstellung nach Schreier (Schreier, 2014) . . . . .	91
5.10. Demographische Daten zu a) Bildungsniveau, b) Social-Media-Nutzung, c) Twitter-Nutzungsdauer und d) Zugangsmethode zu Twitter für die gesamte Stichprobe der Twitter-Umfrage . . . . .	95
5.11. Visualisierung des Auswertungsprozesses wie er sowohl bei den Favorisierungsgründen als auch bei den Wiederfindensgründen (vgl. Abschnitt 5.8) Anwendung fand. Die einzelnen Schritte entsprechen den Schritten des Werkzeugkastenmodells nach Schreier (vgl. Abbildung 5.9). . . . .	97

5.12. Aufteilung der Codes nach den Phasen des Offenen (5.12a) und Axialen (5.12b) Kodierens . . . . .	98
5.13. Häufigkeit des Favorisierens, Entfavorisierens und Favoritenlistenbesuchs im Überblick	99
5.14. Die sechs Reaktionsmöglichkeiten, die das klassische Like abgelöst haben (Screenshot von: <a href="http://newsroom.fb.com/news/2016/02/reactions-now-available-globally/">http://newsroom.fb.com/news/2016/02/reactions-now-available-globally/</a> ) . . . . .	109
5.15. Signifikanzprüfung durch visuelle Begutachtung der Konfidenzintervalle . . . . .	111
5.16. Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Aufbewahrensstrategien . . . . .	112
5.17. Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Aufbewahrensstrategien gruppiert nach Bewertung der Aufbewahrensfunktionen . . . . .	113
5.18. Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien . . . . .	114
5.19. Stärke der Ablehnung bzw. Zustimmung zur Verwendung von Wiederfindensstrategien gruppiert nach Häufigkeit des Aufbewahrens . . . . .	115
5.20. Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien gruppiert nach Einschätzung des Wiederfindensprozesses . . . . .	116
5.21. Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien gruppiert nach Bewertung der Twitter-Optionen für Wiederfinden . . . . .	117
5.22. Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien gruppiert nach Frustrationserfahrung . . . . .	118
5.23. Korrelationsmatrix für Häufigkeits- und Bewertungsvariablen zu Aufbewahren und Wiederfinden auf Twitter . . . . .	119
5.24. Kodierschema (5.24a) der Wiederfindensgründe mit Häufigkeit der Kodennutzung (5.24b) . . . . .	129
6.1. Verortung unterschiedlicher Logdatenerhebungsmethoden in einem Koordinatensystem mit den Achsen <i>Detailgrad der Logdaten</i> und <i>Entwicklungsaufwand</i> . . . . .	144
6.2. Architektur der Chrome-Erweiterung zur Sammlung der Clickstream-Daten . . . . .	152
6.3. Überblick über die Interaktionsmöglichkeiten mit einem Tweet und die zugehörigen CSS-Klassen . . . . .	153
6.4. Schematische Darstellung der detaillierten Ereignisverarbeitung . . . . .	154
6.5. Entity-Relationship-Modell der Datenbank zur Speicherung der Interaktionsevents . . . . .	155
6.6. Die Pop-ups der beiden Experience-Sampling-Varianten in Gegenüberstellung. . . . .	159
6.7. Sessionspezifische Nutzerfeatures mit unterschiedlichen Inaktivitätsgrenzen im Vergleich . . . . .	163
6.8. Visuelle Zusammenfassung der wichtigsten Vorverarbeitungsschritte . . . . .	164
6.9. Boxplot (a) und Tabelle (b) zur Darstellung der Verteilung der Studienteilnahmedauer in Tagen . . . . .	164
6.10. Boxplot (a) und Tabelle (b) der Anzahl Ereignisse pro Studienteilnehmer . . . . .	165
6.11. Twitter-Nutzung der Studienteilnehmer im Tagesverlauf . . . . .	167
6.12. Charakterisierung der Studienteilnehmer in Bezug auf Sessioncharakteristika . . . . .	168
6.13. Anzahl an Sessions pro Nutzer im Verhältnis zu dessen durchschnittlicher Anzahl an Events pro Session . . . . .	169
6.14. Anteil der Ereigniskategorien an den Gesamtereignissen kategorisiert in Gruppen nach Sessiondauer . . . . .	170
6.15. Anteil der Ereigniskategorien an den Gesamtereignissen zu festen Tageszeitpunkten	171
6.16. Veröffentlichungs- und Betrachtungszeitpunkte der gesammelten Tweets . . . . .	172
6.17. Der allererste Tweet gesendet von Jack Dorsey am 21. März 2006 ) (Screenshot von: <a href="https://twitter.com/jack/status/20">https://twitter.com/jack/status/20</a> ) . . . . .	173

6.18. Veröffentlichungs- und Betrachtungszeitpunkte der Tweets zweier Nutzer im Vergleich. Die Grafiken zeigen unterschiedliches Nutzungsverhalten. . . . .	174
6.19. Platzhaltermaße ( <i>Proxies</i> ) verortet auf einem Kontinuum ihrer Ausrichtung von sehr liberal bis sehr konservativ . . . . .	176
6.20. Die drei vorgestellten <i>Proxies</i> Tweetalter, <i>Hover</i> — <i>Hover</i> und <i>Klick</i> — <i>Klick</i> und deren Konsequenz für die Parameter: <i>Anzahl Wiederfindenssitzungen</i> , <i>Anzahl wiederfundener Tweets</i> , <i>Anzahl wiederfindender Nutzer</i> . . . . .	178
6.21. Boxplot (a) und Tabelle (b) der Zeitperiode (Lebenszyklus) zwischen erstem und erneutem Klick . . . . .	179
6.22. Exemplarische Darstellung von Histogramm und Quantil-Quantil-Plot für die Variable <i>DURATION</i> zur visuellen Prüfung auf Normalverteilung . . . . .	182
6.23. Durchschnittliche Differenz der Mittelwerte der beiden Gruppen mit 95% Konfidenzintervall für die Variablen <i>DURATION</i> und <i>SEARCHEVENTCOUNT</i> . Die rotgestrichelte Linie zeigt den Nulleffekt an, der für einen signifikanten Unterschied nicht berührt werden dürfte. . . . .	183
6.24. Vergleich der Re-finding-Strategien bei kurzen und langen Wiederfindensperioden. Als Proxy für die verwendete Strategie dient die jeweilige Twitter Unterseite. . . . .	184
6.25. Nutzerprofilzusammenfassung für den Account @Livefyre. Die beteiligten Links der in der Tabelle aufgeführten Bi-Gramme (2.-5.) sind farblich hervorgehoben. Top 5 Bi-Gramm-Ereignisse, die vor einem Nutzerprofilbesuch auftreten. PSV steht für <i>profile summary viewed</i> und ist das HTML-Overlay, welches die Zusammenfassung des Nutzerprofils zeigt. . . . .	185
6.26. Verteilung der Re-finding-Gründe für die 276 Re-finding-Aktionen . . . . .	186
6.27. Visualisierung der Zusammensetzung erfolgreicher, erfolgloser und normaler Sessions . . . . .	187
6.28. Die vier unterschiedlichen Ergebnisszenarien bei der Analyse erfolglosen Verhaltens veranschaulicht anhand von vier verschiedenen Variablen . . . . .	188
6.29. Aufteilung der Tweets in die unterschiedlichen Inhaltskategorien als Ergebnis des Kodiervorgangs . . . . .	193
6.30. Tabellarische und visuelle Darstellung (ROC-Kurven) der Klassifikationsergebnisse der 15 unterschiedlichen Modelle . . . . .	197
6.31. Die Top 25 Merkmale basierend auf dem Maß <i>mean decrease accuracy</i> des Modells, welches auf Tweet- und Sessionmerkmale basiert. Die Merkmale sind nach Gruppenzugehörigkeit unterschiedlich farblich hervorgehoben. . . . .	198
6.32. Klassifikationsleistung (Accuracy) im Verhältnis zur Anzahl genutzter Variablen. Mit drei Top-Features kann ein identisches Ergebnis erreicht werden wie mit 80 Merkmalen oder mehr. . . . .	199
8.1. Visualisierung von Konfiguration und Anzeige der Webbenachrichtigung favorisierter Tweets (Modifikation der Screenshots von: <a href="https://twitter.com/settings/web_notifications">https://twitter.com/settings/web_notifications</a> bzw. <a href="https://twitter.com">https://twitter.com</a> ) . . . . .	217
8.2. Vergleich zweier Interleaving-Methoden. Grün hervorgehobene Tweets sind ältere, als potentielle Wiederfindensziele identifizierte Tweets, die in den aktuellen <i>Stream</i> integriert werden. (Modifikation eines Screenshots von: <a href="https://twitter.com">https://twitter.com</a> ) . . . . .	219
8.3. Erweiterte Suchfunktion der Twitter-Suche, die das Durchsuchen der gesendeten Tweets der Follower erlaubt (Screenshot von: <a href="https://twitter.com/search-home">https://twitter.com/search-home</a> ) . . . . .	221
8.4. Exemplarische Darstellung eines hervorgehobenen Wiederfindensziels (Modifikation eines Screenshots von: <a href="https://twitter.com">https://twitter.com</a> ) . . . . .	222

9.1. Visualisierung der Instrumentalisierung der neuen Reply-Funktion zur Erstellung eines Threads an zusammenhängenden Tweets

(Screenshot von: <https://twitter.com/larakaaa>) . . . . . 231

# Tabellenverzeichnis

3.1. Suche in bibliographischen Datenbanken nach dem Term Twitter (Stand: November 2016) . . . . .	42
5.1. Demographische Daten zu Alter und Geschlecht der beiden Stichproben im Vergleich	94
5.2. Kategoriensystem/Kodierschema der Favorisierungsgründe als Resultat der qualitativen Datenauswertung . . . . .	100
5.3. Interpretation von logistischen Regressionskoeffizienten bei unterschiedlichen Referenzskalen . . . . .	121
5.4. Die 41 Merkmale, die zur Vorhersage des Frustrationsniveaus genutzt werden. Features mit * besitzen einen normalen und logarithmierten Wert . . . . .	123
5.5. Koeffizienten des logistischen Regressionsmodells . . . . .	124
5.6. Überblick über Maße der Anpassungsgüte sowie die Klassifikationsleistung des Modells . . . . .	125
5.7. Paarweise Berechnung des Kappa-Koeffizienten für alle Codes des Kategoriensystems . . . . .	127
5.8. Paarweise Berechnung des Kappa-Koeffizienten auf der Basis der Oberkategorien . . . . .	128
6.1. Suche in bibliographischen Datenbanken nach Literatur zum Thema Logdatenanalyse . . . . .	142
6.2. Twitter-Account-Informationen der Studienteilnehmer . . . . .	166
6.3. Vergleich der Häufigkeit, mit der Tweets mit und ohne spezieller Interaktion eine erneute Interaktion erfahren . . . . .	177
6.4. Verteilung von Tweets mit und ohne speziellen Interaktion über kurze und lange Wiederfindensperioden . . . . .	180
6.5. Verteilung von Tweets mit und ohne @Mention über kurze und lange Wiederfindensperioden . . . . .	181
6.6. Verortung aller Sessionvariablen in die vier möglichen Ergebnisszenarien als Ergebnis des Bootstrapping-Vergleichs . . . . .	189
6.7. Überblick über die Features, die für die Vorhersage genutzt wurden, klassifiziert in Feature-Gruppen. Features, die mit einem † gekennzeichnet sind, haben absolute und relative Werte. Features, die mit einem * versehen sind, besitzen einen Maximal- und einen Durchschnittswert. . . . .	196
B.1. Verteilungstest der 58 Session-Variablen mit Hilfe von Shapiro-Wilk. . . . .	X





# Abkürzungsverzeichnis

<b>API</b>	Application Programming Interface
<b>AUC</b>	Area under the curve
<b>CSS</b>	Cascading Style Sheets
<b>DOM</b>	Document Object Model
<b>ESM</b>	Experience Sampling-Methode
<b>FAV</b>	Favourite on Twitter
<b>GT</b>	Grounded Theory
<b>HTML</b>	Hypertext Markup Language
<b>IR</b>	Information Retrieval
<b>ISV</b>	Informationssuchverhalten
<b>IVF</b>	Informationsverhaltensforschung
<b>LIS</b>	Library and Information Science
<b>MM</b>	Mixed Methods
<b>MUR</b>	Message Uncertainty Ratio
<b>OOB</b>	Out of Bag
<b>ORM</b>	Object Relational Mapping
<b>OSN</b>	Online Social Network
<b>POC</b>	Part of Conversation
<b>PDO</b>	PHP Data Objects
<b>PIM</b>	Personal Information Management
<b>qIA</b>	qualitative Inhaltsanalyse
<b>ROC</b>	Receiver Operating Characteristic
<b>RT</b>	Retweet on Twitter
<b>SERP</b>	Search Engine Result Page
<b>SVM</b>	Support Vector Machine
<b>TLA</b>	Transaction Log Analysis
<b>UI</b>	User Interface
<b>URL</b>	Uniform Resource Locator



# 1

## Einleitung

*Social Bots*, *Fake News* und *Filter Bubble*. Diese Schlagwörter werden in den Medien häufig in einem Atemzug mit Social-Media-Anwendungen genannt. Social-Media-Anwendungen und deren Auswirkungen auf unsere alltägliche Lebenswelt stehen spätestens seit der tragenden Rolle, die sie bei politischen Entscheidungen im Jahr 2016 eingenommen haben — wie dem Referendum zum Austritt Großbritanniens aus der EU (Brexit) oder den Präsidentschaftswahlen in den USA — im Fokus der medialen Öffentlichkeit. In beiden Fällen haben Plattformen wie Facebook oder Twitter erwiesenermaßen starken Einfluss auf die Meinungsbildung und somit die Wahlentscheidung genommen (Howard & Kollanyi, 2016; Kollanyi et al., 2016).

Social-Media-Nutzung ist mittlerweile vollkommen in unseren Alltag und bisweilen auch in unsere Arbeitswelt integriert. 62% aller US-Amerikaner beziehen ihre Nachrichten über Social-Media-Anwendungen wie Facebook oder Twitter (Gottfried & Shearer, 2016). Studien zur Nutzung von Social Media in der Versicherungsbranche können zeigen, dass die Nutzung von Social-Media-Plattformen während der Arbeitszeit stark positiv mit der Leistung von Arbeitnehmern korreliert (Leftheriotis & Giannakos, 2014). Social-Media-Anwendungen werden also nicht nur in der Freizeit genutzt, sondern sind auch dabei, sich in der Arbeitswelt zu integrieren. Dies zeigen auch Studien zu unternehmensinternen Social-Media-Anwendungen, die die Kommunikation zwischen Mitarbeitern im Unternehmen verbessern sollen (DiMicco et al., 2008). Es drängt sich die Frage auf: Welche Konsequenzen ergeben sich aus der Tatsache, dass sich immer mehr Menschen sehr häufig oder fast ausschließlich über Social-Media-Plattformen informieren? Social Media durchdringt alle Lebensbereiche und stellt uns vor politische und gesellschaftliche Herausforderungen. Nicht nur die Gesellschaft ist durch die Omnipräsenz der Social-Media-Einflüsse vor Herausforderungen gestellt. Jeder Einzelne kann bei der Nutzung von Social-Media-Plattformen mit Problemen konfrontiert werden. Ein Problem stellt der Umgang mit Tweets aus der Perspektive des *Personal Information Management* (PIM) dar. Als PIM bezeichnet man sowohl alle Aktivitäten, die beim Umgang mit persönlichen Informationen auftreten können — wie das Finden, das Organisieren und das Verwalten von Informationen — als auch die wissenschaftliche Disziplin, welche dieses Verhalten untersucht. Die Herausforderungen, die im Kontext von PIM und Social-Media-Anwendungen entstehen können, bilden die Motivation der Arbeit, welche im Folgenden dargelegt wird.

## 1.1. Motivation und Zielsetzung

Die Motivation der Arbeit ergibt sich aus drei Beobachtungen:

1. Social-Media-Anwendungen wie Twitter dienen nicht nur der Aufrechterhaltung von Freundschaften, als Erinnerungsfunktion für Geburtstage oder zum Austausch der neuesten Gerüchte. Sie werden auch, wie bereits oben beschrieben, zur Informationssuche und für seriöse Anwendungszwecke eingesetzt. Im Zuge von Exploratory-Search-Szenarien wie sie von Marchionini beschrieben werden — also bei länger andauernden teilweise auch komplexen Informationsbedürfnissen, wie der Buchung einer Reise oder dem Kauf eines neuen Entertainment-Systems — werden nicht nur klassische Informationsquellen verwendet, sondern es wird immer häufiger auch auf Social Media als Informationsquellen zurückgegriffen (A. T. Chen, 2015; Pálsdóttir, 2014; Sin & Kim, 2013). Eine der von Case und Given formulierten „Ten New Lessons of Information Behaviour Research“ argumentiert, dass formale Informationsquellen und rationales Suchverhalten menschliches Informationssuchverhalten nur zu einem gewissen Grad charakterisiert (Case & Given, 2016, S.361). Stattdessen führen sie eine Reihe an empirischen Studien an, in denen evident wird, dass informelle Quellen wie Social Media häufig als Informationsquellen für alltägliche Informationsbedürfnisse in Bereichen wie Gesundheit, Familie und Finanzen dienen oder auch in Krisensituationen wie Naturkatastrophen konsultiert werden (Yates & Partridge, 2015; Murthy & A.Longwell, 2013).
2. *Personal Information Management* (PIM) und die Probleme, die sich mit der alltäglichen Verwaltung von persönlicher Information ergeben, sind nicht resistent gegenüber technologischem Wandel. Bergman und Whittaker argumentieren: „PIM problems are deep-rooted and general, rather than dependent on a specific technology or type of data“ (Bergman & Whittaker, 2016, S.30). Mit Social-Media-Anwendungen als Informationsquellen ergeben sich also auch im Kontext von PIM und Re-finding neue Probleme und Herausforderungen, die identifiziert, beschrieben und gelöst werden müssen. Denn gutes PIM ist ein wichtiger Bestandteil zur Bewältigung des Alltags. So schreiben Jones und Bruce: „For each of us individuals, better PIM means a better use of our precious resources (time, money, energy, attention) and, ultimately, a better quality of our lives“ (Jones & Bruce, 2005, S.1).
3. Twitter per se macht es aufgrund seines Designs Nutzern nicht einfach, Informationen effektiv und effizient zu verarbeiten. Gomez-Rodriguez, Gummadi und Schölkopf können mithilfe einer großangelegten, quantitativen Studie zeigen, dass Nutzer, die mehr als 30 Tweets pro Stunde erhalten — was etwa 100 *Followees*<sup>1</sup> entspricht — Anzeichen einer Informationsüberflutung zeigen und nicht mehr alle Tweets entsprechend verarbeiten können (Gomez-Rodriguez et al., 2014). Folgende weitere Charakteristika können die Informationsverarbeitung auf Twitter durchaus schwierig gestalten, woraus sich Probleme für den PIM-Kontext ergeben:
  - Tweets sind kurz und können Anhänge enthalten, die, um ein volles Verständnis des Tweets zu bekommen, geöffnet werden müssen.
  - Der Strom an Tweets, also die eigene *Timeline*, besteht aus Tweets von mehreren Absendern gleichzeitig und besitzt außer der umgekehrt-chronologischen Anordnung (neueste Tweets zuerst) keine weitere Organisation.
  - Die Bestimmung des Grads an Interessantheit eines Tweets — man kann dies auch als Relevanzbewertung bezeichnen — muss oft schnell, in einem kleinen Zeitfenster er-

---

<sup>1</sup>Accounts denen ein Nutzer folgt. Antonym zu Follower, Accounts, die einem Nutzer folgen.

folgen, was sich auf die Art der Bewertung auswirken kann (Counts & Fisher, 2011). Erschwert wird diese Bewertung zusätzlich, wenn es sich um Informationen handelt, die im aktuellen Kontext nicht von unmittelbarer Relevanz sind. Die Bewertung/Einschätzung, ob eine Information für die Zukunft von Relevanz sein könnte, ist ohnehin mit einem sehr hohen kognitiven Aufwand verbunden und fehleranfällig (Bruce, 2005). Wenn Nutzer schließlich eine negative Relevanzentscheidung treffen und sich mit dem Problem konfrontiert sehen, zu verpasster Information zurückzukehren zu müssen, ist dies oft ein komplexer Suchvorgang. Auf Twitter verstärkt sich die Komplexität dieser Handlung aufgrund der oben angeführten Charakteristika und der Tatsache, dass Tweets sehr schnell die *Timeline* verlassen können, also sich der Kontext, in dem ein Tweet wahrgenommen wird, rasch ändert.

- Twitter besitzt den Charakter eines Echtzeitinformationsnetzwerkes, das einen Fokus auf aktuelle Ereignisse legt, deren Inhalte aber eher kurzlebig, manchmal auch oberflächlich und somit vermeintlich nicht von längerer Relevanz sind. Jack Dorsey, der Gründer von Twitter, bekräftigt dieses Bild, indem er sich in einem Interview zur Plattform wie folgt äußert: „I don’t go back in time. You’re kind of as good as your last update“ (Dorsey nach Sarno (2009a)).

Zusammenfassend zeigen die Beobachtungen, dass Social-Media-Anwendungen — und Twitter im Speziellen — immer mehr an Relevanz für die Informationssuche gewinnen, sich aber Nutzer dieser Plattformen auch mit PIM-Problemen konfrontiert sehen. Diese Probleme potenzieren sich in diesem Kontext, da diese Plattformen als Echtzeitdienste gedacht und konzipiert sind, die eine längere Relevanz von Informationen in ihrem Design nicht vorsehen. Für Twitter stellt die Schwierigkeit der Informationsverarbeitung per se ein Problem dar. Dies hat vor allem auch Auswirkungen auf die Langzeitnutzung von Informationen und PIM-Verhalten auf Twitter.

Ziel der Arbeit ist es, die Disziplin des *Personal Information Management* zu benutzen, um mit ihr auf das Informationsverhalten von Twitter-Nutzern zu blicken. Im Speziellen wird deren Umgang mit bereits gesichteten Tweets betrachtet. Das übergeordnete Ziel dieser Promotionsarbeit ist es, zu erforschen, wie häufig Twitter-Nutzer zu bereits gesehener Information zurückkehren, welche Motive sie haben, dies zu tun und wie sie dabei vorgehen. Im Fokus steht die Beobachtung, Dokumentation, Beschreibung und Interpretation von Verhalten. Dies geschieht unter Verwendung eines Sequential-Mixed-Methods-Ansatzes, der — aufbauend auf unterschiedlichen Methoden der Datenerhebung und Auswertung — das Problem sowohl auf qualitative und subjektive als auch auf quantitative und damit objektive Weise betrachtet. Konkret werden zwei Studien durchgeführt und präsentiert: eine Umfrage sowie eine Logdatenstudie. In einem nächsten Schritt werden auch Lösungsansätze und Designvorschläge erarbeitet, die das beobachtete Verhalten unterstützen können und somit zu einer besseren Social Media bzw. Twitter *User Experience* beitragen sollen.

## 1.2. Arbeitsdefinitionen

Vor weiteren Ausführungen werden die beiden wesentlichen Begriffe der Arbeit *Personal Information Management* und Social Media definiert.

### **Definition *Personal Information Management***

Jones und Teevan definieren PIM wie folgt:

„Personal information management or PIM is both the practice and the study of the activities people perform to acquire, organize, maintain, retrieve, use and control the distribution of information items such as documents (paper-based and digital), Web pages, and email messages for everyday use to complete tasks

(work-related and not) and to fulfill a person's various roles (as parent, employee, friend, member of community, etc.)“ (Jones & Teevan, 2007a, S.3).

Zwei Aspekte der Definition sind für den vorliegenden Kontext besonders von Interesse: Erstens die Tatsache, dass generische Informationseinheiten das Ziel von PIM-Verhalten sein können, was somit auch Tweets impliziert. Zweitens bezieht sich PIM-Verhalten sowohl auf konkrete Arbeitsaufgaben als auch auf Alltagssituationen, die Informationen erfordern, um die Bedürfnisse von Personen zu erfüllen. Eine ausführliche Diskussion von PIM-Definitionen und das Verhältnis des PIM zu Informationsverhalten im Allgemeinen besitzt, wird in Abschnitt 2.1 gegeben. Dort wird auch die Frage thematisiert, inwiefern Tweets persönliche Informationen darstellen und als Untersuchungsobjekt der PIM-Forschung gelten können. Im PIM-Kontext werden Sammlungen von Informationsobjekten einer Person, wie E-Mails oder Papierdokumente, als einzelne *personal information collections* (PIC) bezeichnet. Mehrere Kollektionen zusammengenommen bilden den *personal space of information* (PSI) einer Person (Jones & Teevan, 2007a, S.10-11). Fokus der Arbeit liegt folglich auf der PIC der Tweets, die eine Person konsumiert.

PIM beschreibt das Problem, dass Informationen heutzutage zwischen verschiedenen Medienarten (analog, digital), verschiedenen Plattformen, verschiedenen Versionen und verschiedenen Endgeräten verstreut sind. Diese Arbeit legt den Fokus auf ein Kernproblem des PIM, das Wiederfinden von Informationen: „Core problems with respect to PIM include the refinding of previously encountered information from shared space as well as search and retrieval of information from personal space“ (Barreau et al., 2008, S.18:1). Eine konkrete Definition von Wiederfinden nehmen Teevan, Capra und Perez-Quinones vor (Teevan, Capra & Perez-Quinones, 2007). Sie definieren Wiederfinden wie folgt: „Re-finding. The process of finding information that has been seen before“ (Teevan, Capra & Perez-Quinones, 2007, S.24). Re-finding oder Wiederfinden beschreibt also den Prozess, bei dem Information, die man bereits einmal gesehen hat, erneut aufgesucht wird. Diese Definition findet sich auf ähnlich auch bei anderen Autoren. Interessant ist dabei, dass selten auf den Zeitabschnitt referenziert wird, der zwischen dem initialen und dem erneuten Auffinden liegt. In dieser Arbeit wird deshalb für eine Differenzierung des Verhaltens entlang verschiedener Dimensionen (z.B. Zeit, Erinnerungsvermögen etc.) plädiert, die Wiederfinden in erneute Aufrufe, (*re-access*) tatsächliches Wiederfinden (*re-finding*) und Wiederentdecken (*re-discovery*) aufteilt.

Die Studien der vorliegenden Arbeit fokussieren auf das Wiederfinden von Tweets vor allem in Bezug auf Tweets aus der *Timeline* des Nutzers, also Tweets, die von anderen Nutzern gesendet wurden. Das Wiederfinden eigener gesendeter Twitter-Nachrichten bleibt dabei nicht unberücksichtigt, wird aber nicht im Detail analysiert.

### Definition Social Media

Den zweiten zentralen Begriff der Arbeit stellt Social Media dar. In der Arbeit wird stets der englische Terminus Social Media in der deutschen Schreibweise verwendet. Auf eine Verwendung der häufig zu lesenden deutschen Übersetzung *soziale Medien* wird verzichtet, da der Begriff einerseits lediglich eine Übersetzung darstellt und als solcher — im Gegensatz zu Social Media — nicht im Duden zu finden ist und andererseits das Adjektiv *sozial* im Deutschen eine umfangreichere Bedeutung besitzt.

Definitionen von Social Media existieren sehr viele, wobei auch unterschiedliche Disziplinen ihre eigenen Standpunkte bei der Diskussion der Bedeutungsbestimmung des Begriffs einnehmen (Geißler, 2010; Kaplan & Haenlein, 2010; Kietzmann et al., 2011; Leonardi et al., 2013). Als Arbeitsdefinition wird die wohl bekannteste, da am häufigsten zitierte, Social-Media-Definition von Kaplan und Haenlein gewählt.<sup>2</sup> Kaplan und Haenlein definieren Social Media unter Verwendung der Konzepte *Web 2.0* und *User Generated Content* wie folgt:

„Social-Media is a group of Internet-based applications that build on the ideological and technological foun-

<sup>2</sup>Der Artikel besitzt 8280 Zitationen auf Google Scholar. Stand: Dezember 2016.

“...and that allow the creation and exchange of User Generated Content“ (Kaplan & Haenlein, 2010, S.61).

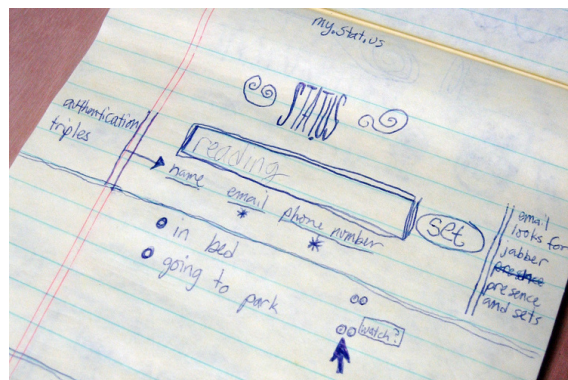
Ferner nehmen sie eine Klassifikation von Social-Media-Applikationen entlang sechs Arten vor, wobei sie für jede Variante eine Beispielanwendung nennen. Diese sechs Arten sind:

1. Social Networks (Facebook)
2. Kollaborative Projekte (Wikipedia)
3. Blogs und Mikro-Blogs (Twitter)
4. *Content Communities* (Youtube, Flickr)
5. Virtuelle Spielewelten (World of Warcraft)
6. Virtuelle Welten (Second Life)

Diese Einteilung ist sicherlich diskutabel und könnte um weitere Arten wie gemeinschaftlich erstellte Nachrichtenseiten (Reddit) oder kollaborative Bookmark-Seiten (Delicious) ergänzt werden. In der vorliegenden Arbeit dient die Mikroblog-Plattform und Social-Media-Anwendung Twitter als Untersuchungsgegenstand. Diese wird im Folgenden näher vorgestellt.

### 1.3. Die Social-Media-Plattform Twitter

Der Abschnitt stellt die Social-Media-Plattform Twitter vor. Hierfür wird ein kurzer historischer Abriss über die Entwicklung Twitters gegeben, wobei ein besonderer Fokus auf die für den PIM-Kontext relevanten Twitter-Funktionen und deren Veränderung über die Zeit gelegt wird. Abschließend werden aktuelle Nutzungszahlen angeführt.



**Abbildung 1.1.:** Erster Sketch des Twitter-Vorgänger Stat.us von Jack Dorsey (Dorsey, 2006)

Als Grundlegung Twitters gilt der erste Tweet, den Gründungsmitglied und Mitentwickler Jack Dorsey am 21. März 2006 sendet. Neben Jack Dorsey sind Biz Stone und Evan Williams maßgeblich an der Entwicklung Twitters beteiligt. In einem Interview mit David Sarno von der *Los Angeles Times* im Jahr 2009 schildert Jack Dorsey die Ideen und Beweggründe, die die Entwicklung Twitters beeinflusst haben (Sarno, 2009b). In diesem Interview nennt Dorsey drei Einflussfaktoren, von

denen Twitter besonders geprägt ist: der Webblog-Software *LiveJournal*<sup>3</sup>, den Echtzeitmöglichkeiten, die *Instant Messenger* bieten, sowie dem ortsunabhängigen Senden von Nachrichten, welches bei SMS-Nachrichten möglich ist. Abbildung 1.1 zeigt den ersten Entwurf Twitters, der bereits im Jahr 2000 entstand. Dem Entwurf ist zu entnehmen, dass Twitter anfangs noch *Stat.us* genannt wurde. Erst im Laufe der Zeit wurde *Stat.us* in Twitter umbenannt. Diesen Schritt erklärt Dorsey wie folgt:

„[W]e looked in the dictionary [...] and we came across the word 'twitter', and it was just perfect. The definition was 'a short burst of inconsequential information' and 'chirps from birds'. And that's exactly what the product was. The whole bird thing: bird chirps sound meaningless to us, but meaning is applied by other birds. The same is true of Twitter: a lot of messages can be seen as completely useless and meaningless, but it's entirely dependent on the recipient“ (Dorsey nach Sarno (2009b)).

Twitter existiert nun seit mehr als 10 Jahren und kann sowohl aus unternehmerischer Perspektive (Börsengang, Übernahme von anderen Unternehmen) als auch in Bezug auf die technologische Entwicklung auf zahlreiche entscheidende, erinnerungswerte Momente zurückblicken. Twitter selbst stellt einige dieser Momente auf ihrer Website vor (Twitter, 2016c). Für den vorliegenden Kontext sind besonders die technologischen Entwicklungen, also die Entwicklung der Plattformfunktionalitäten, von Bedeutung. Einige dieser Veränderungen werden herausgegriffen und dargestellt. Es ist zu beachten, dass hier auch neuere Veränderungen vorgestellt werden, die zum Zeitpunkt der Studien noch nicht umgesetzt waren und dementsprechend keinen Einfluss auf die Formulierung von Leitfragen und die Konzeption der Studien genommen haben.

- **November 2006:** Twitter erhält ein großes Update, das u.a. die Favoriten-Funktion, Direktnachrichten und die Twitter-Suche umfasst (Stone, 2006). Die Favoriten-Funktion ermöglicht es Nutzern, Tweets hervorzuheben, indem sie mit einem Stern markiert werden. Die markierten Nachrichten werden in einer separaten Liste gespeichert. Direktnachrichten sind Nachrichten, die nicht öffentlich sichtbar sind sondern direkt an Twitter-Nutzer gesendet werden. Die Suche ermöglicht das Durchsuchen aller Twitter-Nachrichten, die zu einem bestimmten Term gepostet wurden.
- **Mai 2007:** Einführung der @Reply- bzw. @Mention-Funktionalität (Stone, 2007a). Twitter reagiert mit der Einführung auf das Verhalten von Nutzern, die damit begonnen hatten, eine Adressierung von Tweets an andere Nutzer via dem @-Symbol vorzunehmen.
- **August 2007:** Chris Messina verwendet zum ersten Mal den Hashtag (vgl. Abbildung 1.2), einem mit einem Doppelkreuz versehenen Term.
- **Januar 2009:** ein US Airways Flugzeug führt eine Notlandung im Hudson River in New York City durch. Das erste Foto vom Unfallort wird über Twitter verbreitet, noch bevor traditionelle Medien davon berichten. Dies ist ein entscheidender Moment, der die Wahrnehmung von Twitter als ein seriöses Echtzeitmedium befeuert.
- **August 2009:** *Project Retweet: Phase One* beginnt und damit erste Tests zur Retweet-Funktion (Stone, 2007b). Twitter greift damit abermals das Verhalten ihrer Nutzer auf, die Tweets von anderen Anwendern kopiert und weitergeleitet haben und diese mit *RT* gekennzeichnet haben. Dies wird als neues Feature in die Plattform integriert. Ende 2009 wird die Funktionalität schließlich plattformübergreifend angeboten.
- **2009:** Twitter führt die Listen-Funktion ein, die es Nutzern erlaubt, ihre Followees zu strukturieren und separate *Timelines* zu erstellen.
- **2011:** Twitter breitet sich immer mehr aus und versucht durch Übernahmen (*TweetDeck*) und Partnerschaften (mit Apple für die Integration von Twitter in iOS) auf immer mehr Endgeräten und Betriebssystemen Fuß zu fassen. In diesem Jahr berichtet Twitter von 1

<sup>3</sup><http://www.livejournal.com/>



Milliarde gesendeten Tweets pro Woche und 100 Millionen aktiven Nutzern pro Monat.

- **2015:** Die Direktnachrichten-Funktion wird stark überarbeitet. Das 140-Zeichen-Limit für diese wird aufgehoben. Zudem können sie nun auch an Nicht-Follower und an mehrere Personen gleichzeitig (Gruppen) gesendet werden.
- **Januar 2015:** Das While-you-were-away-Feature, das die Top-Tweets vergangener Tage bzw. Stunden nach einem längeren Nicht-Besuch an den Kopf der *Timeline* stellt, wird eingeführt (Rosania, 2015).
- **April 2015:** Die RT-Funktion wird überarbeitet und ermöglicht fortan die Variante *retweet with comment*, so dass nicht nur das Weiterleiten eines Tweets möglich ist, sondern dieser auch in Form eines Kommentares in eigene Tweets eingebettet werden kann (Shu, 2015).
- **November 2015:** Eines der ältesten Features, die Favoriten-Funktion, die vormals durch einen Stern visualisiert war, wird nun in *Hearts* umbenannt und durch ein Herz visualisiert. Dies stellt die Like-Bedeutung der Funktion in den Vordergrund (Kumar, 2015). Diese Umstellung wird später noch näher diskutiert.
- **Mai 2016:** Die 140-Zeichen-Beschränkung, die seit Twitters Gründung bestand hat, wird gelockert, indem einige Tweet-Inhalte wie *@Mentions*, URLs oder Medien-Anhänge (Bilder, Videos) nicht mehr zum 140-Zeichen-Limit gerechnet werden. In einem späteren Update wird es sogar möglich sein eine *@Reply*, also eine Antwort, an Nutzer zu schreiben, ohne die *@Mention*-Syntax zu verwenden, sowie eigene Tweets weiterzuleiten (Sherman, 2016).
- **Dezember 2016:** Die Suchfunktion wird stark überarbeitet. Tweets der SERP werden nicht mehr in umgekehrter chronologischer Reihenfolge angezeigt, sondern nach Relevanz sortiert (L. Huang, 2016).

Im Folgenden werden die Twitter-Funktionen, die für das PIM-Verhalten und das Wiederfinden von Bedeutung sind — was in Abschnitt 3.2 anhand von relevanten Studien zu diesen Themen noch weitaus detaillierter beschrieben wird — vorgestellt.

### Die Hashtag-Funktion

Der Hashtag, also die Verbindung aus dem Schriftzeichen Doppelkreuz (#) und einem Schlüsselwort, besitzt seine Wurzeln im *Internet Relay Chat* (IRC) und wurde dort zur Kennzeichnung von Raumnamen genutzt. Chris Messina benutzte den Hashtag auf Twitter erstmalig im August 2007 und wird als Urvater des Hashtags bezeichnet (vgl. Abbildung 1.2).



**Abbildung 1.2.:** Erstmalige Verwendung des Hashtags auf Twitter im August 2007 durch den Nutzer Chris Messina (Screenshot von: <https://twitter.com/chrismessina/status/223115412>)

ps://twitter.com/chrismessina/status/223115412)

Erst im Juli 2009 implementiert Twitter die Hashtag-Funktionalität wie sie bis dato existiert und definiert einen Hashtag wie folgt:

„A hashtag — written with a # symbol — is used to index keywords or topics on Twitter. This function was created on Twitter, and allows people to easily follow topics they are interested in. [...] People use the hashtag symbol (#) before a relevant keyword or phrase in their Tweet to categorize those Tweets and help them show more easily in Twitter search“ (Twitter, 2016e).

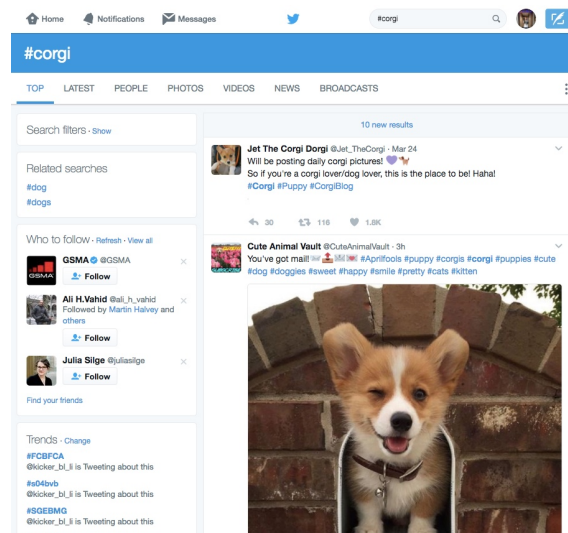
Der Hashtag als Markierungselement bietet die Möglichkeit der Informationserschließung, um Themen zu identifizieren, Tweets zu diesen Themen durch Verwendung des jeweiligen Hashtags zuzuordnen und schließlich bei einem Klick auf einen Hashtag ein Suchergebnis zu erhalten, das sämtliche Tweets mit diesem Hashtag enthält. Es handelt sich um eine weitere Form des *social tagging*, das seit der Ankunft des partizipativen Web 2.0 einen Höhepunkt um das Jahr 2010 hatte und auch intensiv Thema der PIM-Forschung war (z. B. in den Arbeiten von Heckner, Heilemann und Wolff (Heckner et al., 2009) oder Körner et al. (Körner et al., 2010)); dessen Nutzung jedoch, wie z. B. beim Social-Bookmarking-Dienst *Delicious*<sup>4</sup>, sich nie wirklich durchgesetzt hat. Der Hashtag stellt ein Instrument zur Informationserschließung dar. Inwiefern der Hashtag für PIM-Verhaltensweisen und Wiederfinden von Tweets genutzt wird, werden die relevante Literatur sowie die eigenen Studien zeigen.

### Die Twitter-Suche

Die Twitter-Suche ist eine Funktion, die seit dem Gründungsjahr Twitters in die Plattform integriert ist. Die Suche nach Tweets hat die Information-Retrieval-Forschungslandschaft nachhaltig geprägt, wie Abschnitt 3.2.1.1 zeigen wird. Twitter beschreibt seine Such-Funktionalität wie folgt:

„There are many ways to use search on Twitter. You can find Tweets from yourself, friends, local businesses, and everyone from well-known entertainers to global political leaders. By searching for topic keywords or hashtags, you can follow ongoing conversations about breaking news or personal interests“ (Twitter, 2016f).

Interessant ist die Tatsache, dass die Suchfunktion der Website andere Funktionalitäten besitzt als die Suchfunktionen der mobilen Twitter-Clients für iOS und Android.



**Abbildung 1.3.:** Screenshot der Twitter-Suche respektive der Twitter-SERP im April 2017

(Screenshot von: <https://twitter.com/search?src=typd&q=%23corgi>)

<sup>4</sup><https://del.icio.us/>

Abbildung 1.3 zeigt das UI der Twitter-Suche. Es wird deutlich, dass Ergebnisse nach den Eigenschaften *top*, also nach Relevanz, und *latest*, also nach Aktualität, sortiert werden können. Ferner können explizit Personen gesucht werden oder Tweets mit besonderen Medien-Anhängen, wie Fotos und Videos. Die Suchleiste verfügt über die Autocomplete- und Autosuggest-Funktion. Zudem besitzt die Twitter-Suche über eine Erweiterte-Suche, die die Verwendung von Suchoperatoren wie *@Mentions* erlaubt, um Tweets eines bestimmten Senders oder Tweets, die diesen Nutzer erwähnen, zu finden. Ferner können nähere zeitliche Einschränkungen oder Ortsbeschränkungen (*near you*) vorgenommen werden. Klassische Boolesche Operatoren wie *AND* und *OR* sind ebenso verwendbar. Auch der Hashtag kann als Operator verstanden werden, da nur Tweets mit diesem Tag zurückgeliefert werden. Suchanfragen können gespeichert und somit durch einfache Klicks wiederholt werden, ohne die Suchanfragen noch einmal explizit eingeben zu müssen, was ein *monitoring* von bestimmten Themen erleichtert. Jeder Account kann bis zu 25 Suchanfragen speichern.

Mit dem zuvor bereits erwähnten Update der Suchfunktion vom Dezember 2016 werden die Tweets der *search engine result page* (SERP) nach einer Suche nicht mehr in umgekehrt-chronologischer Reihenfolge angegeben, sondern sind nach Relevanz sortiert. Lisa Huang, eine Softwareentwicklerin bei Twitter, die für den Bereich *search quality* zuständig ist, begründet die Umstellung wie folgt: „Twitter is live and real-time, so naturally Twitter search must surface recent, yet still relevant, results. Historically, these search results have been largely presented in reverse chronological order. In many cases however, the most recent results may not be what the searchers are looking for“ (L. Huang, 2016). Neben den Hauptproblemen, wie der richtigen Balance aus Tweet- und Account-Ergebnissen oder Tweets mit und ohne Medien-Anhängen in der SERP der Twitter-Suche, beschreibt Huang die positiven Effekte dieses neuen Rankings, nämlich, dass Nutzer, die mit der neuen Twitter-Suche konfrontiert werden, nicht nur mehr mit der Twitter-SERP interagieren, sondern auch mehr Tweets verfassen und mehr Zeit auf Twitter verbringen (L. Huang, 2016).

### Die Retweet-Funktion

Die Einführung der RT-Funktion ist, wie bereits beschrieben, eine Reaktion auf das Verhalten der Nutzer. Vor der offiziellen Einführung des Features haben diese den Tweet-Text anderer Nutzer kopiert, dies mit *RT: [@Mention des Autors]*, *retweeting [@Mention des Autors]* oder *via [@Mention des Autors]* gekennzeichnet und an die eigenen *Follower* gesendet (boyd et al., 2010). Dieses Verhalten beschreibt die erste Phase der RT-Funktionalität. Die zweite Phase ist durch die Einführung des RT-Buttons gekennzeichnet, welcher zum Ende des Jahres 2009 für sämtliche Nutzer zugänglich wurde. Die dritte Phase der RT-Funktion wurde durch die Überarbeitung im Jahr 2015 eingeführt. Sie erlaubt das Hinzufügen von Kommentaren. Insgesamt hat die RT-Funktion also eine Veränderung in drei Phasen durchlaufen.

Die aktuelle Variante der RT-Funktion definiert Twitter wie folgt: „A Tweet that you share publicly with your followers is known as a Retweet“ (Twitter, 2016a). Dort wird ferner auch die Kommentarfunktion erläutert, die es erlaubt, dass Tweets nicht nur weitergeleitet werden können, sondern auch als Anhang für einen eigenen Tweet dienen können und somit der eigene Text als Kommentar zum Tweet gesehen werden kann. In der Beschreibung der Funktionalität wird auch explizit auf das Weiterleiten eigener Tweets eingegangen, eine Funktionalität die erst Mitte 2016 (siehe oben) eingeführt wurde. Explizit heißt es dort: „In addition to sharing other’s Tweets on Twitter, you can Retweet or Quote Tweet your own Tweet. This function is particularly useful when you want to re-post one of your older Tweets because it is relevant again“ (Twitter, 2016a). Das Weiterleiten von älteren Tweets setzt implizit auch deren Wiederfinden voraus. Zu Beginn der vorliegenden Forschungsarbeit war diese Funktionalität noch nicht vorhanden, so dass keine Analysen zu dieser Perspektive durchgeführt wurden. Die Analyse von eigenen Tweets, die von einem selbst weitergeleitet wurden, erlaubt einen ganz neuen Blick auf Wiederfinden auf Twitter und stellt eine Möglichkeit für zukünftige Studien dar. Unabhängig von dieser Perspektive ist die RT-Funktion

für die Frage nach Wiederfindensverhalten von Relevanz, wie Abschnitt 3.2.2.2 zeigen wird. Abschließend bleibt zu erwähnen, dass Tweets auch via Direktnachricht oder als Link geteilt werden können.

### Die Favoriten-Funktion

Die Favoriten-Funktion wurde im November 2006 im Zuge eines der ersten großen Updates als neue Funktion in Form eines klickbaren Sterns vorgestellt (Stone, 2006). Durch das Klicken eines Sterns können sowohl Tweets von anderen als auch eigene Tweets in die Favoritenliste aufgenommen werden. Die Favoritenliste ist über die Profilseite eines Nutzers erreichbar und kann öffentlich von allen Nutzern eingesehen werden. Obwohl die Funktion seit Beginn Teil der Twitter-Plattform ist, wurde diese nur wenig beworben und es etablierten sich auch nur wenige eindeutige Nutzungspraktiken (Paßmann & Gerlitz, 2014).

Im November 2015 wurde der Stern schließlich durch ein Herz ersetzt und zugleich die Favoriten-Funktion in eine Like-Funktion geändert. Twitter kündigte die Änderung wie folgt an:

„We are changing our star icon for favorites to a heart and we'll be calling them likes. We want to make Twitter easier and more rewarding to use, and we know that at times the star could be confusing, especially to newcomers. You might like a lot of things, but not everything can be your favorite“ Kumar (2015).

Begründet wurde der Wechsel zu einem anderen Icon und einem anderen mentalen Konzept mit dem Argument, dass das Herz ein universales Symbol sei, das über zahlreiche Kulturen, Sprachen und Zeitzonen konsistent interpretiert werden würde, sowie eine stärkere Ausdruckskraft besäße, mit dem mehrere Emotionen kommuniziert werden könnten (Kumar, 2015). Bei den Nutzern jedoch wurde diese Veränderung nicht sehr positiv aufgenommen. Trotz dieser anfänglichen Ressentiments seitens der Nutzer hat sich die Like-Funktion mittlerweile etabliert. Bis auf einen unterschiedlichen Button (Herz statt Stern) hat sich das Nutzungserlebnis wenig verändert. *Gelikte* Tweets werden immer noch im eigenen Profil in einer separaten Liste geführt, wie dies für die Favoriten seit jeher Standard war und können dort vom Nutzern wieder aufgefunden werden. Ob und in welchem Ausmaß die Favoriten- bzw. Like-Funktion eine Rolle im PIM-Verhalten auf Twitter spielt, stellt einen großen Untersuchungspunkt der Arbeit dar (vgl. Abschnitt 5.6). Da die Datenerhebung für beide Studien noch vor der Funktionsänderung endete, wird in der Arbeit auch stets die Bezeichnung Favoriten-Funktion genutzt, da die Like-Funktion noch nicht existierte.

Es wird deutlich, dass die Analyse einer Social-Media-Plattform wie Twitter den Forschenden durchaus vor Probleme stellen kann, da unvermittelte Veränderungen auftreten können, die Forschungsfragen oder gestaltete Experimente obsolet werden lassen. Erfreulicherweise wurden während des Datenerhebungszeitraums der beiden Studien für die vorliegende Arbeit keine maßgeblichen Veränderungen an den für das PIM-Verhalten relevanten Twitter-Funktionen durchgeführt. Diese Diskussion wird mit den kritischen Anmerkungen zur Arbeit (vgl. Abschnitt 9.2.2) nochmals aufgegriffen.

### Twitter-Nutzungszahlen

Laut einem aktuellen Bericht des *Pew Research Center* vom November 2016 nutzen 21% aller US-Bürgerinnen und Bürger Twitter (Greenwood et al., 2016, S.3). Im Vergleich zu anderen Social-Media-Plattformen wie Facebook, Pinterest oder Instagram besitzt Twitter somit die geringsten Nutzerzahlen (Greenwood et al., 2016). In einer Prognose von 2014 sieht das Statistik-Portal *Statista*<sup>5</sup> die Anzahl der monatlich aktiven Nutzer in Deutschland bei etwa 5,7 Mio. Im Vergleich zu 2014 würde dies einen Anstieg um 1,4 Mio. Nutzer bedeuten, was bei einer Einwohnerzahl von 82 Mio. etwa 7% der Bevölkerung entspricht (Statista, 2014). Offizielle Nutzerzahlen bietet die Twitter-Website, mit der Angabe von 313 Mio. aktiver Nutzer im Monat. Von diesen sind 79% aller Accounts außerhalb der Vereinigten Staaten registriert (Twitter, 2016d). Twitter wird vor allem von jüngeren Nutzern aus der Altersgruppe der 18- bis 29-Jährigen (36%) genutzt. Bei der

<sup>5</sup><https://de.statista.com>

Altersgruppe 65+ liegt der Prozentsatz der Nutzer nur bei 10% (Greenwood et al., 2016, S.6). Ferner stammen die Nutzer der Plattform überwiegend aus einer gebildeten Bevölkerungsschicht. 29% der Nutzer verfügen mindestens über einen College-Abschluss. Hinsichtlich des Geschlechts und geographischer Regionen (Stadt/Land) lassen sich keine Unterschiede in den Nutzerzahlen feststellen (Greenwood et al., 2016, S.6). Beinahe alle Twitter-Nutzer (93%), besitzen zusätzlich auch einen Facebook-Account. Die Nutzung Twitters erfolgt vorwiegend über mobile Endgeräte, wobei die Angabe „82% Active users on mobile“ auch so interpretiert werden könnte, dass 82% der Nutzer neben dem Web-Client den Service auch über einen mobilen Client nutzen (Twitter, 2016d).

## 1.4. Gliederung der Arbeit

### Kapitel 1: Einleitung

Das Kapitel führt in das Thema ein, legt Motivation sowie Zielsetzung der Arbeit dar und definiert die beiden wichtigen Konzepte *Personal Information Management* und Social Media. Abschließend wird der Untersuchungsgegenstand, die Social-Media-Anwendung Twitter, näher vorgestellt.

### Kapitel 2: Personal Information Management

Kapitel 2 stellt die Disziplin des *Personal Information Management* näher vor. Es besteht aus zwei Teilen. Im ersten Teil wird das Verhältnis zwischen allgemeinem Informationsverhalten und PIM vorgestellt. Im Zuge dessen wird die Rolle von Social Media als Untersuchungsgegenstand der Informationsverhaltensforschung angesprochen und darauf Bezug genommen, inwiefern Twitter bzw. Tweets persönliche Informationen darstellen, bzw. welche Informationscharakteristika Tweets besitzen. Im zweiten Teil wird PIM als akademische Disziplin vorgestellt, wobei ein Überblick über relevante Forschungsarbeiten erfolgt. Ein besonderer Fokus liegt auf der Darstellung von Forschungsarbeiten zum Wiederfindensverhalten in unterschiedlichen Kontexten wie mit E-Mails und Webseiten.

### Kapitel 3: Social-Media-Forschung: Twitter und Social-Media-PIM im Fokus

Dieses Kapitel stellt die Dimensionen der Social-Media-Forschung am Beispiel von Twitter vor. Einleitend werden unterschiedliche Perspektiven, aus denen Twitter untersucht wird, vorgestellt. Ein besonderer Schwerpunkt liegt zum einen auf Social-Media-Forschungsarbeiten, die Bezug zu *Personal Information Management* im Allgemeinen besitzen sowie zum anderen auf Studien, die Twitter aus der Perspektive der PIM-Praktiken untersuchen. Besonders die Abschnitte zur Verwaltung und Wiederfinden von Social-Media-Inhalten geben den für die vorliegende Promotionsarbeit relevanten aktuellen Forschungsstand wieder.

### Kapitel 4: Forschungsdesign

Das Forschungsdesign der Promotionsarbeit wird in Kapitel 4 erläutert. Zunächst wird eine Einführung in mögliche Forschungsstrategien wie Triangulation und Mixed-Methods-Studien gegeben. Anschließend wird das Sequential-Mixed-Design — eine Variante eines Mixed-Methods-Ansatzes — als Forschungsdesign vorgestellt, das der vorliegenden Promotionsarbeit zu Grunde liegt. Abschließend werden die konzeptionellen Gedanken der beiden Studien der Arbeit — die Twitter-Umfrage sowie die Twitter-Logstudie — erörtert.

### Kapitel 5: Studie I: Twitter-Umfrage

Kapitel 5 stellt die erste der beiden großen Studien der Arbeit, die Twitter-Umfrage, vor. Kapitel 5 und 6 haben dabei stets einen ähnlichen Aufbau. Zu Beginn werden die Leitfragen der jeweiligen Studie formuliert. Anschließend wird die verwendete Datenerhebungsmethodik sowie das Studiendesign näher beschrieben. Als Rekrutierungsstrategie für die Twitter-Umfrage wurde unter anderem ein Crowdsourcing-Panel verwendet, um eine möglichst heterogene Stichprobe

an Twitter-Nutzern ansprechen zu können. Diese Form der Probandenrekrutierung wird näher präsentiert. Aufgrund des hohen Anteils an verbalisierten, qualitativen Daten, die während der Twitter-Umfrage erhoben werden, besitzt Kapitel 5 eine ausführliche Diskussion von qualitativen Auswertungsmethoden und stellt die beiden Methoden Grounded Theory sowie die qualitative Inhaltsanalyse näher vor. Anschließend werden die drei Schwerpunkte der Twitter-Umfrage und deren Ergebnisse vorgestellt. Diese sind: die Rolle der Favoriten-Funktion für das PIM-Verhalten, die quantitativen Erkenntnisse über Aufbewahrens- und Wiederfindensverhalten, sowie die Analyse von Wiederfindensmotiven.

#### **Kapitel 6: Studie II: Twitter-Logstudie**

Kapitel 6 beschreibt die zweite große Studie der Arbeit, die Twitter-Logstudie. Einen großen Raum nimmt die Darstellung der Logdatenanalyse als Methode ein, die sowohl aus informationswissenschaftlicher als auch aus methodologischer Perspektive vorgestellt wird. Im Anschluss daran wird die eigens für die Studie konzipierte und implementierte *Chrome* Webbrowser-Erweiterung, die zur Sammlung der Logdaten genutzt wurde, vorgestellt. Die Logstudie wurde zusätzlich um ein *Experience Sampling* ergänzt. Diese Methodik wird ebenfalls erläutert. Vor der Darstellung der quantitativen Beschreibung des Wiederfindensverhalten in all seinen Facetten wird Twitter-Verhalten im Allgemeinen beschrieben, da die Art der Daten und die Aussagen, die aus ihnen abgeleitet werden können, ebenfalls ein Novum im Kontext der Twitter-Forschung bilden. Wiederfindensverhalten wird anschließend aus mehreren Gesichtspunkten empirisch beleuchtet. Ein Machine-Learning-Experiment, das es sich zum Ziel setzt, Wiederfindensziele (Tweets) vorherzusagen, stellt den Schlußpunkt des Kapitels dar.

#### **Kapitel 7: Forschungsfragen erneut betrachtet: Erkenntnisse aus Literatur und Studien vereint**

Dieses Kapitel nimmt eine Meta-Inferenz im Sinne des Forschungsdesigns vor und diskutiert die Ergebnisse der beiden Studien vor dem Hintergrund der bestehenden Erkenntnisse aus der Literatur. Dies wird besonders unter Berücksichtigung der beiden PIM-Praktiken Aufbewahren und Wiederfinden durchgeführt. Das Kapitel schließt mit der Identifikation von Problemen, die Nutzer beim Wiederfinden von Tweets auf Twitter besitzen. Im Folgekapitel werden diese Schwierigkeiten aufgegriffen und als Anschlusspunkte zur Vorstellung von Lösungsmöglichkeiten bzw. Designvorschlägen genommen werden, um Wiederfindensverhalten zu unterstützen.

#### **Kapitel 8: Wiederfindensverhalten unterstützen: Lösungsansätze und Designvorschläge**

Wie bereits erwähnt, greift Kapitel 8 die in der Diskussion erkannten Nutzerprobleme beim Wiederfinden von Tweets auf und stellt mögliche Lösungsansätze in Form von Designvorschlägen vor. Konkret werden vier Varianten vorgestellt, die Assistenz beim Wiederfindensverhalten bieten, wobei explizit die beiden Verhaltensweisen, die in Umfrage und Logstudie als Wiederfindensstrategien erkannt wurden, als Ansatzpunkte für Lösungsvorschläge genutzt werden.

#### **Kapitel 9: Fazit und Ausblick**

In einem Fazit werden die wesentlichen Erkenntnisse der Arbeit nochmals zusammengefasst und kritische Anmerkungen zur Arbeit, wie die Schwierigkeiten, die eine längere Untersuchung einer Social-Media-Anwendung wie Twitter in sich birgt, thematisiert. Die Promotionsarbeit schließt mit einem Ausblick auf zukünftige Studien, die aus den offenen Fragen der vorliegenden Studien resultieren.

# 2

## Personal Information Management

Kapitel 2 beginnt in Abschnitt 2.1 mit einer Argumentation für die Informationsverhaltensforschung als zentrales informationswissenschaftliches Forschungsparadigma. Darauf aufbauend wird der Zusammenhang zwischen Informationsverhalten und PIM herausgearbeitet. Abschnitt 2.1.1 wirft einen Blick auf Social Media aus der Perspektive der Informationsverhaltensforschung (IVF) und zeigt mögliche Fragestellungen bzw. Herausforderungen auf. Die Abschnitte 2.1.2 sowie 2.1.3 dienen als Argumentationsgrundlage für Twitter und Tweets als Untersuchungsgegenstände der PIM-Forschung. PIM als Forschungsfeld und seine Forschungsgebiete werden in Abschnitt 2.2.1 vorgestellt. Im Anschluss daran werden die wesentlichen PIM-Aktivitäten im Detail charakterisiert. Die Aktivitäten Aufbewahren (vgl. Abschnitt 2.2.2) und Organisieren/Verwalten (vgl. Abschnitt 2.2.3) werden dabei weniger umfangreich behandelt. Der Fokus liegt auf Abschnitt 2.2.4, also der Darstellung des Wiederfindensverhaltens. Da in Abschnitt 2.1.3 argumentiert wird, dass Tweets Charakteristika von E-Mails und Webdokumenten in sich vereinen, wird explizit auf das Wiederfindensverhalten im Umgang mit E-Mail und Webdokumenten fokussiert. Abschnitt 2.3 fasst die wesentlichen Erkenntnisse und Charakteristika der PIM-Praktiken zusammen und verdeutlicht, welche möglichen Forschungsfragen sich für das PIM-Verhalten bzw. Wiederfindensverhalten auf Twitter ergeben.

Einige Inhalte dieses Kapitels wurden vorab als Artikel im Themenheft Informationsverhalten der Zeitschrift *Information, Wissenschaft und Praxis* veröffentlicht (Meier, 2015). Dies betrifft insbesondere die Abschnitte 2.1.1 sowie Abschnitt 3.2.1.1 des Folgekapitels.

### 2.1. Vom Informationsverhalten zur Verwaltung persönlicher Information

Zu Beginn dieses Abschnitts steht das Verhältnis der Informationswissenschaft zum Informationsverhalten im Fokus. Darauf aufbauend wird herausgearbeitet, wie PIM als wesentlicher Bestandteil alltäglichen Informationsverhaltens gesehen werden kann.

Die Informationswissenschaft ist sowohl thematisch als auch methodisch von einer starken Heterogenität geprägt, welche sich unter anderem aus ihrer Nähe zu benachbarten Disziplinen (z. B. Psychologie, Sprachwissenschaft, Informatik, etc.) bzw. Interdisziplinarität ableiten lässt (Sarace-

vic, 2009). Zins stellt fest, dass es keine einheitliche Auffassung davon gibt, was Informationswissenschaft ausmacht, so dass die Disziplin zahlreiche unterschiedliche Traditionen und Forschungsansätze besitzt (Zins, 2007, S.335). Als Beispiel führt er das klassische Gegensatzpaar „objective approaches versus cognitive approaches“ (Zins, 2007, S.335) an, welches dem über viele Jahre lange dominierenden systemorientierten Ansatz den von Ingwersen und Järvelin als *cognitive turn* (Ingwersen & Järvelin, 2005) formulierten kognitiven Standpunkt gegenüberstellt, der den Nutzer und dessen Kontext in den Fokus rückt. Betrachtet man die von Zins mithilfe der Delphi-Methode von Informationswissenschaftlern gesammelten Definitionen des Faches, fällt auf, dass nur drei Personen den Term *behaviour* verwenden, nur drei Definitionen den Term *user* enthalten und nur ein Wissenschaftler mit dem Term *interaction* arbeitet (Zins, 2007). Dahingegen enthalten sieben Definitionen den Term *systems* und fünf Definitionen den Term *retrieval* (Zins, 2007). Es lässt sich innerhalb der Definitionen also durchaus ein systemorientierter Schwerpunkt erkennen. Dies ist nicht verwunderlich, folgt man den Ausführungen von Fisher, Erdelez und McKechnie, die erklären, dass sich die Beobachtung und Analyse des Informationsverhaltens als Subdisziplin der *Library and Information Science* erst Mitte der 90er Jahre herausbildet und der Begriff des *Information Behaviour* sich erst zu dieser Zeit zu etablieren beginnt (K. Fisher et al., 2005, S.XIX). Es lassen sich aber schon früher Spuren in informationswissenschaftlicher Forschung erkennen, die belegen, dass die Beobachtung und Analyse des Informationsverhaltens eine zentrale methodische Herangehensweise in der Informationswissenschaft darstellt.

Schon in den frühen 1960er Jahren, bei der Evaluation von IR-Systemen auf der Basis von Logdaten, standen nicht nur Systemparameter im Fokus, sondern auch das Verhalten des Nutzers wurde thematisiert (Peters, 1993). Dieser Umstand wird in Abschnitt 6.2.3 noch weiter vertieft. Zahlreiche Forscher argumentieren, dass das Informationsverhalten stets ein zentrales Forschungsparadigma der Informationswissenschaft war und ist (Robertson & Hancock-Beaulieu, 1992; Bates, 1999; Wilson, 2000; Saracevic, 2009). Bates nennt drei zentrale Fragestellungen für informationswissenschaftliche Forschung (Bates, 1999). Die von ihr als soziale Frage formulierte Fragestellung zielt auf Informationsverhalten: „How do people relate to, seek, and use information?“ (Bates, 1999, S.1048). Wilson argumentiert, dass verhaltensorientierte Forschung eine starke Tradition in der Informationswissenschaft besitzt. Er untermauert dies, indem er Umfragen zur Bibliotheksnutzung (*The Library Survey*) aus den 1920er Jahren als Beispiele anführt (Wilson, 2000, S.50). Saracevic resümiert in einem Artikel zur Geschichte der Informationswissenschaft: „[I]nformation science has two orientations: one that deals with information retrieval techniques and systems and the other that deals with [...] human information behaviour“ (Saracevic, 2009, S.2585). Elsweiler und Hammwöhner sehen „[e]ine alle relevanten Informationsprozesse umfassende Theorie des Informationsverhaltens“ (Elsweiler & Hammwöhner, 2015, S.4) auf der Basis von empirischen Studien als ein Desiderat informationswissenschaftlicher Forschung und charakterisieren „[e]inen auf Informationsverhalten bzw. informationelle Handlung fokussierten Forschungsansatz der Informationswissenschaft“ (Elsweiler & Hammwöhner, 2015, S.7) als denjenigen Forschungsansatz, der es ermöglicht, einen Rahmen für die vorhandenen, jedoch bisher unverbundenen Forschungsparadigmen in der Informationswissenschaft zu schaffen (Elsweiler & Hammwöhner, 2015, S.7). Zudem plädieren sie für einen eher weitgefassten Informationsverhaltensbegriff, der nicht nur bewusstes und aktives Handeln (bzw. Interaktion) mit Information unter Informationsverhalten fasst, sondern auch unbewusstes Handeln, wie die Blickbewegungen des Auges oder Informationsvermeidung, als Informationsverhalten interpretiert und dabei auch auf die Erfassung kognitiver Zustände abzielt (Elsweiler & Hammwöhner, 2015). So ist ein derartiger Informationsverhaltensbegriff kein Novum in der Informationsverhaltensforschung, betrachtet man frühere Definitionen.

Wilson definiert *Information Behaviour* als „the totality of human behavior in relation to sources and channels of information, including both active and passive information seeking, and information use“ (Wilson, 2000, S.49). Er zählt explizit Verhaltensweisen, wie Konversationen zwischen



Personen oder das Betrachten von Fernsehwerbung, zum Informationsverhalten. Zudem beklagt er, dass in der Wissenschaft der Fokus auf dem *Information Seeking Behaviour*, also der Suche nach Information unter der Verwendung von Informationssystemen wie Bibliothekskataloge oder Web-suchmaschinen, liegt (Wilson, 2000, S.49). Man muss Wilsons Beobachtung sicherlich zustimmen, aber auch anerkennen, dass die Studien zum *Information Seeking* zu zahlreichen Informationsverhaltensmodellen, wie z. B. von Ellis (Ellis, 1989), Kulthau (Kuhlthau, 1991), oder Marchionini (Marchionini, 1995) geführt haben, die noch heute von Bedeutung sind und das Gebiet nachhaltig beeinflusst haben.<sup>1</sup> Case und Given definieren *Information Behaviour* als:

„[IB] encompasses information seeking as well as the totality of other unintentional or serendipitous behaviours (such as glimpsing or encountering information), as well as purposive behaviors that do not involve seeking, such as actively avoiding information. The term also includes the broader context of how individuals 'deal with' information in their lives [...]“ (Case & Given, 2016, S.6).

Auch bei Case und Given wird nochmals deutlich: Informationsverhalten ist nicht nur Informationssuche. Auch die Begegnung mit Information oder das Vermeiden von Information gehören zum Informationsverhalten. Informationsverhalten benötigt keine konkrete Aufgabe (*task*) auch keine arbeitsbezogene Aufgabe (*work task*) und betrifft alltägliche Handlungen, wie den Kauf eines Autos oder die Planung einer Reise (Case & Given, 2016, S.6-7).

Ausgehend von diesen Definitionen, die Informationsverhalten als den allumfassenden Umgang mit Information charakterisieren, würde man ohne weiteres davon ausgehen — obwohl dies in den Definitionen nicht explizit erwähnt wird — dass die Verwaltung (das Management) persönlicher Informationen ein wesentlicher Bestandteil des alltäglichen Informationsverhaltens ist. Betrachtet man jedoch relevante Literatur, so finden sich beispielsweise in der zweiten Ausgabe von *Looking for Information* keine Studien mit PIM-Bezug (Case, 2007). Erst in der vierten und neuesten Ausgabe wird eine PIM-Arbeit von Whittaker et al. zitiert und folglich findet auch der Begriff *Personal Information Management* Verwendung (Case & Given, 2016). Eine Darstellung von dem was PIM ausmacht, fehlt in beiden Ausgaben jedoch völlig. Eine Definition von Informationsverhalten, welche PIM-Bezug besitzt, wird von Fisher, Erdelez und McKechnie formuliert: „[W]e conceptualize information behavior as including how people need, seek, manage, give, and use information in different contexts“ (K. Fisher et al., 2005, S.XIX). Diese Auffassung von Informationsverhalten schlägt sich auch inhaltlich in Fisher, Erdelez und McKechnie's Buch wieder, da hier Arbeiten mit PIM-Bezug, wie die PAIN-Hypothese von Bruce, ausführlich vorgestellt werden (K. Fisher et al., 2005). Dass PIM Bestandteil des Informationsverhaltens ist und somit eine Subdisziplin der Informationsverhaltensforschung darstellt, scheint Commonsense zu sein, wird aber in der Literatur wenig expliziert. In Abschnitt 1.2 der Einleitung wurde bereits eine Arbeitsdefinition von PIM gegeben. Im Folgenden soll zusätzlich zur Definition eine detaillierte Charakterisierung von PIM vorgenommen werden. Zunächst sollen jedoch folgende Fragen geklärt werden: Was macht Social Media zu einem geeigneten Untersuchungsgegenstand der Informationsverhaltensforschung? Inwiefern handelt es sich bei Tweets um persönliche Information? Welche Informationscharakteristika — im Sinne Bergmans und Whittakers — besitzen Tweets (Bergman & Whittaker, 2016, S.24-26)?

### 2.1.1. Ein Blick auf Social Media aus der Perspektive der Informationsverhaltensforschung

Aus der Perspektive der Informationsverhaltensforschung stellen Social-Media-Applikationen interessante Untersuchungsgegenstände dar. Dies liegt insbesondere daran, dass in Alltagssituationen,

<sup>1</sup>Jede der hier genannten Veröffentlichungen besitzt über 1000 Zitationen in Google Scholar.

die Informationsbedürfnisse hervorrufen, Nutzer zur Tilgung dieser Bedürfnisse informellen Informationsquellen den Vorzug vor formalen Ressourcen geben (Case & Given, 2016, S.361). Studien von Savolainen, Johnson sowie Agosto und Hughes-Hassell zeigen, dass Personen, die nach Informationen, Ratschlägen oder Empfehlungen suchen, zunächst ihre eigenen sozialen Kreise konsultieren, bevor sie eine formale Suche beginnen (Savolainen, 1995; Agosto & Hughes-Hassell, 2005). Johnson argumentiert, dass Informationskanäle dann besonders häufig genutzt werden, wenn sie dem Nutzer den Eindruck vermitteln, Face-to-Face-Kommunikation zu ermöglichen (Case & Given, 2016, S.159). Social-Media-Applikationen besitzen genau diese Unvermitteltheit in der Kommunikation und sind somit der ideale Ort, um das eigene soziale Netzwerk nach Ratschlägen oder Empfehlungen zu bitten. Dass dieses Verhalten zum Alltag auf Social-Media-Applikationen gehört, wird Abschnitt 3.2.1.2 zeigen, welcher Analysen vorstellen wird in denen die Nutzung Twitters als soziales Frage-Antwort-Systeme thematisiert wird. Es zeigt sich, dass Social-Media-Applikationen ideale Untersuchungsgegenstände für die IVF bieten. Khoo fasst die Stärken von Social-Media-Applikationen als Informationsquelle aus mehreren Studien zusammen und listet mögliche Fragestellungen auf, die damit einhergehen (Khoo, 2014). Folgende Punkte nennt Khoo als Stärken von Social Media als Informationsquelle (Khoo, 2014, S.82):

- Bestimmte Informationsarten (Empfehlungen, Ratschläge) können nur von Personen gegeben werden. Zudem können diese eine optimale Kontextualisierung und Personalisierung durchführen.
- Persönlichen Kontakten wird ein hohes Maß an Vertrauen entgegengebracht, weshalb Informationen aus dem eigenen sozialen Netzwerk als glaubwürdiger erscheinen können. Zudem kann auch die Emotionalität der Kommunikation eine Rolle spielen.
- Personen können als Medium agieren und Informationen recherchieren, synthetisieren, aufbereiten oder an andere Personengruppen mit höherer Expertise weiterleiten.

Für Khoo ist die wesentliche Frage, wie sich Social Media als Informationsquellen von anderen, klassischen Quellen unterscheiden und wie dies das Informationssuchverhalten der Nutzer beeinflusst. Er nennt einige konkrete Fragestellungen, von denen einige auch im vorliegenden Kontext von Relevanz sind (Khoo, 2014):

- Welche Parallelen existieren zwischen klassischem Suchverhalten (z. B. mit Websuchmaschinen) zum Suchverhalten mit Social-Media-Applikationen?
- Welche Informationsbedürfnisse, Aufgaben und Gründe bewegen Nutzer dazu, sich an Social-Media-Applikationen zu wenden, bzw. diese vor klassischen Informationsquellen zu präferieren?
- Welche Rolle spielen bestimmte Verhaltensweisen wie Navigation, Suche und *serendipity*?
- Welche Unterschiede lassen sich zwischen verschiedenen Nutzergruppen identifizieren?
- Welche Methoden eignen sich zur Analyse des Informationssuchverhaltens auf Social-Media-Plattformen?

Letztendlich nennt Khoo aber auch Probleme bzw. Herausforderungen, die mit Social-Media-Plattformen und ihrer Untersuchung verbunden sind (Khoo, 2014, S.90):

- Wie im Rahmen der Einleitung bereits skizziert wurde, unterliegen Social-Media-Plattformen einem steten Wandel, was auch bereits in der Einleitung am Beispiel Twitters deutlich wurde. Ständig kommen neue Funktionalitäten hinzu oder werden modifiziert. Neben etablierten Applikationen treten auch immer wieder neue Plattformen auf, die die Social-Media-Ökologie der Nutzer beeinflussen und auch deren Verhalten im Falle der Adaption neuer Plattformen beeinflussen.
- Unterschiedliche Social-Media-Applikationen ermöglichen die Suche nach Information auf

ganz unterschiedliche Weise. Auch Plattformen gleicher Kategorie, z. B. Microblog-Plattformen wie Tumblr und Twitter, weisen zahlreiche Unterschiede in ihrer Verwendung auf, wodurch sich die Frage aufdrängt, wie diese Unterschiede die Untersuchung des Verhaltens beeinflusst.

Social Media als Untersuchungsgegenstand der IVF beginnt sich gerade erst zu formieren und ist ein sehr junges Feld. PIM als Teilgebiet der IVF ist dahingehend noch beinahe unerforscht. Die vorliegende Arbeit, welche den Umgang mit persönlichen Informationen im Kontext von Social Media in den Fokus rückt, bearbeitet ein kaum erforschtes Gebiet. Inwiefern Tweets als persönliche Information wahrgenommen werden können, wird im Folgenden diskutiert.

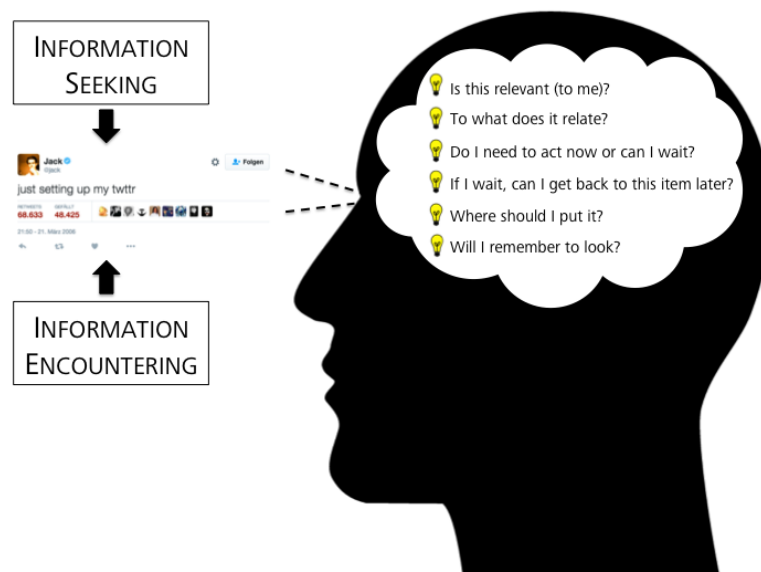
### 2.1.2. Tweets als persönliche Information

Für PIM ist die Frage von Bedeutung, wann Information zu persönlicher Information wird. Besonders im vorliegenden Kontext ist dies durchaus eine berechtigte Frage, die zu klären ist. Ist sämtliche Information in Form von Tweets, die durch die *Twitter-Timeline* läuft, persönliche Information? Ein Argument, das hier angeführt werden kann, ist, dass auf Social-Media-Plattformen — auf Twitter im Besonderen — die Informationen, die man erhält, schon dadurch persönlich werden, da man selbst die Entscheidung darüber trifft, aus welchen Accounts sich der Kreis derjenigen zusammensetzt, denen man folgt und deren Inhalte und Informationen man entsprechend konsumiert. In diesem Sinne muss man aber eher von personalisierter als von persönlicher Information sprechen. Im Folgenden werden drei Argumente angeführt, die ein Plädoyer dafür sein sollen, warum Informationen auf Twitter persönliche Information darstellen. Jones (Jones, 2012, S.23) wie auch Jones und Teevan (Jones & Teevan, 2007a, S.9-10) führen sechs Charakteristika an, die Information zu persönlicher Information werden lassen, indem sie die Relation der Information zum eigenen Selbst darstellen:

- Information gehört mir oder wird von mir kontrolliert.
- Information bezieht sich auf mich.<sup>2</sup>
- Information ist an mich gerichtet.
- Information wird von mir zur Verfügung gestellt/gesendet.
- Information wird/wurde schon einmal von mir wahrgenommen.
- Information ist relevant/nützlich für mich (Aufgabenkontext).

All diese Charakteristika, die Information zu persönlicher Information werden lassen, können auch auf Tweets zutreffen. Dies ist jedoch nicht das einzige Argument. Im Folgenden wird zusätzlich argumentiert, dass es weniger auf die Herkunft bzw. die Art der Information ankommt als auf die kognitiven Prozesse — umgangssprachlich die Gedanken, die sich eine Person über Information macht — welche beim Rezipienten ablaufen. Bergman und Whittaker charakterisieren Information dann als persönlich bzw. als Teil einer persönlichen Informationskollektion, wenn diese als wertvoll oder nützlich für zukünftige Situationen beurteilt wird (Bergman & Whittaker, 2016, S.1). Diese Charakterisierung kann um die Situation ergänzt werden, in der sich Personen an Information erinnern, der für ein vorliegendes Problem Relevanz zugesprochen wird, wobei in einer früheren Begegnung zwischen Nutzer und Information, diese nicht als relevant erachtet und möglicherweise übergangen wurde. Das Entscheidungsproblem, mit dem Nutzer bei der Rezeption von Information jedes Mal konfrontiert werden — also der Frage, ob Information möglicherweise zukünftigen Wert

<sup>2</sup>Im Kontext der immer bedeutsamer werdenden Diskussion um die eigene Privatsphäre beginnen Nutzer im Web nach sich selbst zu suchen, um festzustellen, welche Informationen dort über sie zu finden sind. Dieses Phänomen, das man auch als *self search* bezeichnet, wird u.a. von Marshall untersucht (Marshall & Lindley, 2014).



**Abbildung 2.1.:** Mentale Fragen, die aus Information persönliche Information werden lassen

besitzt und deswegen aufbewahrt werden müsste — bezeichnet Bruce als das Problem des *personal, anticipated information need* (PAIN) (Bruce, 2005). Die Überlegungen bzw. Gedanken Information könnte in Zukunft relevant sein und wäre jetzt relevant, also eine Art retrospektive Relevanz (hervorgerufen durch die Erinnerung an gesehene Information), sind folglich zwei Seiten einer Medaille, die als persönliche Information bezeichnet werden kann. In beiden Situationen versucht der Nutzer ein *mapping* zwischen Informationsbedürfnis und Information herzustellen (Jones & Teevan, 2007a, S.14). Bergman und Whittaker charakterisieren PIM in diesem Kontext auch als Dialog, den wir mit unserem zukünftigen Selbst führen, welches uns mitteilt, welche Informationsbedürfnisse wir in Zukunft haben und uns bei der Entscheidung hilft, ob Information aufbewahrt werden sollte (Bergman & Whittaker, 2016, S.17-18). In diesem Sinne kann auch ein Tweet, an den man sich erinnert ihn in der eigenen *Timeline* gelesen zu haben, persönliche Information sein.

Unabhängig von der medialen Form einer Information — und deswegen auch auf Tweets zutreffend — sind Nutzer bzw. Personen immer mit einer Kette von Entscheidungen konfrontiert, welche in Abbildung 2.1 visualisiert sind. Sobald diese Prozesse angestoßen werden und sobald Information bewertet wird, wird sie zu persönlicher Information.

### 2.1.3. Informationscharakteristika eines Tweets

Neben der Frage, inwiefern Tweets als persönliche Information gelten können, stellt sich zusätzlich die Frage, welche Eigenschaften Tweets als Informationseinheit besitzen. Dabei geht es nicht darum, welche Arten von Information (Text, Bild, Video, Link etc.) in Tweets auftreten, sondern um die *information properties* aus Bergmanns und Whittakers Perspektive (Bergman & Whittaker, 2016, S.24). Da Informationscharakteristika Einfluss darauf haben, ob und wie Informationen aufbewahrt und organisiert werden, ist diese Frage auch in Bezug auf Tweets zu klären. Bergman und Whittaker charakterisieren Informationseinheiten entlang der Dimensionen

- Informationstyp (informativ/handlungsbedürftig)

- Einzigartigkeit der Information
- Art der Ansammlung (passiv/aktiv)

Der Informationstyp (informativ vs. handlungsbedürftig) beschreibt den Umstand, dass ein Informationsobjekt rein informativ sein kann, also vom Nutzer lediglich rezipiert wird, aber auch eine Handlung seitens des Rezipienten verlangt sein kann, wie dies häufig bei E-Mails der Fall ist, bei denen eine Antwort des Nutzers — oft innerhalb eines gewissen Zeitfensters — erwartet wird. Da diese Handlungen meist nicht instantan, also beim direkten Empfang oder Erhalt einer Information durchgeführt werden können, muss der Nutzer Strategien entwickeln, welche ihn an die noch auszuführende Handlung erinnern. Denkbar sind To-do-Ordner oder das Zurücklassen einer E-Mail im Posteingang (Bergman & Whittaker, 2016, S.24).

Die Einzigartigkeit einer Information bzw. Informationseinheit ist charakterisiert durch die Anzahl an möglichen Zugangswege und der Art der Zugangswege (öffentlich vs. privat). Einzigartigkeit bezieht sich also auf die Einfachheit des Zugangs und der Häufigkeit der Zugangswege. Zum Beispiel kann Webinformation auf mehreren Webservern verteilt liegen und durch eine Suchanfrage in einer Websuchmaschine auch mehrfach gefunden werden. Einzigartigkeit definiert sich ferner durch die Frage nach dem Autor, also ob Information von einem selbst erstellt wurde und damit möglicherweise nur am Erstellungsort vorhanden ist. Schließlich ist Einzigartigkeit aber auch persönlich definiert. Bestimmte Informationen, die für die eine Person besonders relevant oder interessant erscheinen und deswegen aufbewahrt und organisiert werden, werden von einer anderen Person als belanglos wahrgenommen (Bergman & Whittaker, 2016, S.26). Einzigartigkeit hat starken Einfluss auf die Aufbewahrensentscheidung, ob und wie Information organisiert wird — z. B. in einer Ordnerstruktur — und folglich auch auf die verwendeten Strategien zum Wiederfinden dieser Information.

Als dritte Eigenschaft von Information — und damit auch Einflussfaktor auf das PIM-Verhalten — kann die Art und Weise gelten, wie sich Information ansammelt. Vergleicht man wiederum E-Mails und Websites, wird deutlich, dass sich E-Mails automatisch im Posteingang ansammeln, ohne dass der Nutzer hier weitere Schritte vornehmen müsste. Websites hingegen müssen aktiv und bewusst aufbewahrt werden z. B. in Form eines Browser-Lesezeichens. Die Frage, wie sich Information ansammelt, hat folglich starken Einfluss auf die Aufbewahrensentscheidung (Bergman & Whittaker, 2016, S.28). Ähnliche Charakterisierungen von Informationstypen findet man bei Cole (Cole, 1982) oder Barreau (Barreau & Nardi, 1995).

Wendet man die hier vorgestellten Eigenschaftsdimensionen von Information auf den Informationstyp Tweet an, kann man feststellen, dass Tweets hauptsächlich als informativ charakterisiert werden können.<sup>3</sup> Tweets können aber auch Handlungsbedarf seitens des Nutzers erfordern, wenn ein Tweet via *@Mention* direkt an den Nutzer gerichtet ist. Des Weiteren können andere Faktoren beim Nutzer Handlungsbedarf erwecken, wie der Wunsch nach Favorisieren oder Weiterleiten, der — ähnlich wie bei einer Antwort auf eine erhaltene E-Mail — nicht immer direkt erfolgen muss, sondern eine gewisse Bedenkzeit erfordern kann, oder erst aufgrund äußerer Reize durchgeführt wird. In Bezug auf die Einzigartigkeit von Tweets muss man feststellen, dass Tweets zwar in den meisten Fällen einzigartig sind, jedoch vor allem bei professionellen Accounts beobachtet werden kann, dass Links zu externen Websites oft häufiger als nur einmal versendet werden, um die Wahrscheinlichkeit zu erhöhen, dass ein Tweet auch von Nutzern aus anderen Ländern und Zeitzonen wahrgenommen werden kann. Natürlich trägt auch die RT-Funktion dazu bei, dass ein Tweet öfter als nur einmal in der *Timeline* eines Nutzers erscheinen kann, sodass auch dadurch ein Tweet an Einzigartigkeit verliert. Tweets sammeln sich grundsätzlich automatisch an, ohne das Zutun des Nutzers. Hypothetisch gesehen könnte man durch ein infinites Scrolling durch die eigene *Timeline*

<sup>3</sup>Ein Großteil der Tweets wird seitens des Nutzers auch einfach übergangen, ohne dass eine konkrete Relevanzentscheidung getroffen wird.

jeden bereits gesendeten Tweet erreichen. Dies wäre aber mit hohem Aufwand verbunden bzw. scheint die Identifikation des gesuchten Tweets ein nahezu unmögliches Unterfangen zu sein, da man bei älteren Tweets vermutlich wenig Erinnerung an Kontextfaktoren — wie Sendezeitpunkt oder Tweets von anderen Nutzern, die diesen Tweet umgeben — als Anker bei der Lokalisierung besitzt. Die Twitter-*Timeline* unterscheidet sich daher auch deswegen von einer E-Mail-Inbox, da Tweets nicht gelöscht werden müssen, sondern automatisch aus dem Wahrnehmungsfeld des Nutzers verschwinden. Um ein persönliches Archiv an relevanten Tweets anzulegen, müssen Nutzer aktiv handeln.

Die Dimensionen der *information properties* zeigen, dass Tweets Charakteristika von E-Mails, Websites aber auch persönlichen Dokumenten vereinen. Bei der Vorstellung der PIM-Verhaltensweisen wird immer wieder auf die Besonderheit von bestimmten Informationstypen Bezug genommen werden.

## 2.2. PIM als Forschungsfeld und sein Forschungsstand

Im Folgenden wird über einen kurzen historischen Abriss von PIM über dessen Definitionen hin zur Vorstellung relevanter PIM-Forschungsarbeiten geleitet. Jones schreibt: „The modern dialog on PIM begins with the publication of Vannevar Bush’s ‘As we may think’“ (Jones, 2012, S.12). Er charakterisiert das von Bush erdachte Memex-System als erstes PIM-Tool, welches dafür konzipiert war, um das menschliche Gedächtnis bei der Verwaltung von Informationen zu unterstützen (Jones, 2012, S.12). Der Begriff *Personal Information Management* wurde zum ersten Mal in den 1980er Jahren gebräuchlich als Reaktion auf die Verbreitung des *personal computer* (PC) und darauf installierte Software (PIM-Tools), die dem Nutzer das Verarbeiten und Verwalten von persönlichen Informationen, via Kontaktbüchern, Terminkalendern oder To-do-Listen, erleichtern (Jones & Teevan, 2007a, S.3). Spuren des PIM können jedoch bereits bis zu den ersten Höhlenmalereien zurückverfolgt werden (Jones & Teevan, 2007a, S.3). In ihrem Buch merken Jones und Teevan an:

„Interest in PIM has increased in recent years; it is now not only a hot topic but also a serious area of inquiry focussing the best work from a diverse set of disciplines including cognitive psychology, human-computer interaction, data-base management, information retrieval, and library and information science“ (Jones & Teevan, 2007a, S.3).

Bezieht man sich auf den zweiten Teil der Aussage, wird deutlich, dass PIM als Forschungsgegenstand von vielen Disziplinen untersucht wird, deren Interessenschwerpunkte durchaus unterschiedlich gelagert sind.

Der erste Teil der Aussage formuliert die These, dass obwohl PIM ein sehr junges Feld ist, PIM-Fragestellungen seitens der Forschung sehr starkes Interesse hervorgerufen haben. Jones und Teevan argumentieren, dass vor allem das Bewusstsein darüber, wie neue Medien und Technologien auch zu Problemen, z. B. dem Informationsüberfluss, führen — da immer mehr Daten, in unterschiedlichen Varianten und Versionen auf den verschiedensten Endgeräten verstreut sind — und die Auseinandersetzung mit diesen Problemen, PIM-Forschung motiviert hat (Jones & Teevan, 2007a, S.4). Gutes PIM ist schließlich für jedes Individuum von essentieller Bedeutung. So schreiben Jones und Bruce: „For each of us individuals, better PIM means a better use of our precious resources (time, money, energy, attention) and, ultimately, a better quality of our lives“ (Jones & Bruce, 2005, S.1). PIM hat also Praxisrelevanz, indem es Einfluss auf unser alltägliches Verhalten nimmt. PIM besitzt natürlich auch wissenschaftliche Relevanz. Bergman und Whittaker argumentieren, dass seitens der Forschung ein zu starker Fokus auf das Informationssuchverhalten neuer Information in großen Informationskollektionen gelegt wird. Ihrer Ansicht nach müsse man aber einen Blick über die Modelle des *information foraging* bzw. *information seeking* hinaus werfen und Modelle und Theorien der Informationsorganisation (*information curation*) erstellen,

nachdem diese unserem alltäglichen Handeln, das vor allem durch Organisation und Navigation in immer größer werdenden eigenen Informationskollektionen geprägt ist, viel mehr entsprechen (Bergman & Whittaker, 2016, S.13). Trotz des gesteigerten Interesses seitens der Wissenschaft am Thema PIM, sowie einer Vielzahl an veröffentlichten Monographien, Konferenzbeiträgen und Journalartikeln, behaupten Bergman und Whittaker, dass überraschend wenig Forschung zum Thema existiert (Bergman & Whittaker, 2016, S.XI). Klar ist, dass die Forschung einen Fokus auf das Informationssuchverhalten legt. Ihre Aussage ist jedoch kaum haltbar, betrachtet man die Anzahl an — auch populärwissenschaftlichen Veröffentlichungen, die als PIM-Ratgeber verstanden werden können — veröffentlichten Monographien zum Thema (Jones & Teevan, 2007b; Jones, 2008, 2012, 2013, 2015; Bergman & Whittaker, 2016).

### **Personal-Information-Management-Definitionen**

So diskussionswürdig wie die Definitionen von Informationswissenschaft und Informationsverhalten sein mögen, umso klarer ist die Auffassung davon, was *Personal Information Management* ausmacht. In Abschnitt 1.2 wurde bereits eine formelle Definition von PIM und seinen Hauptaktivitäten gegeben. Vergleicht man diese Definition mit einer Begriffsbestimmung von Bergman und Whittaker, wird deutlich, dass eigentlich keine Unterschiede in der Auffassung bestehen, was PIM charakterisiert:

„Personal information management (PIM) is the process by which individuals curate their personal data in order to reaccess that data later. Curation involves three distinct processes: how we make decisions about what personal information to keep, how we organize that kept data, and the strategies by which we access it later“ (Bergman & Whittaker, 2016, S.1).

Bergman und Whittaker führen das Drei-Phasen-Modell der Informationspflege (*three-stage model of the curation life cycle*) ein, bestehend aus den Tätigkeiten: Aufbewahren (*keeping*), Verwalten (*management*) und Wiederverwerten (*exploitation*) (Bergman & Whittaker, 2016, S.15). Sie weisen aber darauf hin, dass dieses Modell ebenfalls auf dem PIM-Framework von Jones und Teevan bzw. Jones basiert (Jones & Teevan, 2007a; Jones, 2012), wodurch abermals deutlich wird, dass durchaus ein Konsens darüber besteht, was PIM charakterisiert.

Obwohl die Informationssuche ein wesentlicher Bestandteil des PIM-Verhaltens ist — das Wiederholen von Schritten, die man beim initialen Finden einer Information gemacht hat, zählt zu einer beliebten Wiederfindensstrategie — wird Informationssuche in der vorliegenden Arbeit lediglich aus der Perspektive des Auffindens von Informationen im Twitter-Kontext (vgl. Abschnitt 3.2.1) vorgestellt und nicht auf der Basis der allgemein zu diesem Thema vorhandenen Literatur. Im Folgenden wird bei der Darstellung relevanter PIM-Literatur und -Studien ein Fokus auf Arbeiten gelegt, die speziell das Wiederfindensverhalten analysieren. Da aber Wiederfinden stark davon beeinflusst ist, ob und wie Information aufbewahrt, verwaltet und organisiert wird, können diese PIM-Praktiken nicht übergangen werden, werden jedoch kürzer behandelt. Bevor nun Untersuchungen und Erkenntnisse zu den einzelnen PIM-Praktiken/Hauptaktivitäten vorgestellt werden, wird ein kurzer Abriss zu Dimensionen der PIM-Forschung gegeben.

### **2.2.1. Dimensionen der PIM-Forschung**

Folgt man Boardman und Sasse, so existieren in der PIM-Forschung zwei Hauptgebiete: Empirische Studien, die das PIM-Verhalten von unterschiedlichen Personengruppen in verschiedenen Situationen des Alltags (Arbeit, Freizeit etc.) charakterisieren, sowie Studien zur Gestaltung und Entwicklung von Software-Prototypen (Boardman & Sasse, 2004, S.583). Neben dieser Einteilung kann auch noch eine Klassifizierung hinsichtlich des untersuchten Informationstyps (E-Mails, Notizen, Webseiten, Patientenakten usw.) oder der verwendeten Methodik vorgenommen werden. Auch methodisch wurden in der PIM-Forschung neue Wege beschritten, insofern der hohe per-

sönliche Kontext von PIM eigene Probleme in sich birgt — klassische experimentelle Situationen besitzen nicht genügend ökologische Validität — welche bei der Erforschung von PIM-Verhalten gelöst werden müssen.

### **Dimension: Empirische Studien**

PIM-Forschung per se ist interdisziplinär angelegt, so dass in den ersten Studien zu PIM ein psychologischer Schwerpunkt zu erkennen ist, der sich vor allem auf das Zusammenspiel von Gedächtnis bzw. Erinnerungsvermögen bei der Verwaltung persönlicher Informationen bezieht (Malone, 1983; Lansdale, 1988). Malone untersucht das Verhalten von Büroangestellten im Umgang mit papierbasierten Dokumenten mithilfe von Interviews und einem Rundgang (*guided tours*) durch das Büro (Malone, 1983). Bei der Auswertung legt er besonderen Fokus auf kognitive Aspekte wie dem Erinnerungsvermögen. Er prägt die sehr bekannte Einteilung von Personen in diejenigen, die *filing* bevorzugen, also die Kategorisierung von Dokumenten in Ordnerstrukturen vornehmen, und Personen, die *piling* bevorzugen, also Nutzer, die Informationen zu Haufen stapeln (Malone, 1983). Dabei stellt er fest, dass das Kategorisieren von Dokumenten mit hohem kognitivem Aufwand verbunden ist und durch das Ablegen von Dokumenten in Ordnern die Erinnerungsfunktion, welche Stapel leisten können, verloren geht (Malone, 1983, S.109-110). Lansdale fokussiert in seiner Studie auf den Einfluss, den der Aufgabenkontext auf die Gedächtnisleistung beim PIM-Verhalten hat (Lansdale, 1988). Auch in neueren Arbeiten wird die Rolle des Erinnerungsvermögens beim PIM-Verhalten untersucht. Elsweiler, Baillie und Ruthven analysieren die Rolle des Erinnerungsvermögens beim Wiederfinden von E-Mails besitzt (Elsweiler et al., 2008). Deng und Kollegen entwickeln einen gedächtnisbasierten Ansatz eines Recall-by-context-Suchanfragemodells (Deng et al., 2011). Aktuelle Forschung im kognitionspsychologischen Kontext begibt sich sogar auf die unterste Ebene der neurologischen Psychologie. Benn et al. untersuchen die Hypothese, dass Nutzer die Navigation in Ordnerstrukturen gegenüber Schlüsselwort-Suche bevorzugen (vgl. Abschnitt 2.2.4), da diese Wiederfindensstrategie Ähnlichkeit zur Navigation bzw. Orientierung in der echten Welt besitzt (Benn et al., 2015). In einem Experiment mit 17 Probanden unter der Zuhilfenahme von funktioneller Magnetresonanztomographie (fMRT) können Benn und Kollegen zeigen, dass bei Navigation in einem Dateisystem tatsächlich dieselben Hirnareale verwendet werden, wie bei einer Navigation im Gelände, was letztendlich die Begründung für die Präferenz von Navigation gegenüber Suche darstellen soll (Benn et al., 2015, S.3). Neben dem Gedächtnis werden auch andere psychologische Einflussfaktoren auf PIM, wie der der Persönlichkeit, thematisiert. Massey et al. untersuchen in zwei Studien, ob und wie sich Persönlichkeitsmerkmale auf das PIM-Verhalten von Nutzern auswirken (Massey et al., 2014). Zur Erhebung der Persönlichkeitsmerkmale verwenden sie das bekannte und weitverbreitete Big-Five-Modell (Massey et al., 2014, S.3698). Sie stellen fest, dass PIM-Verhalten nicht rein funktional oder rational ist, sondern stark von der eigenen Persönlichkeit beeinflusst ist. Beispielsweise konnte bei Personen mit neurotischen Persönlichkeitszügen gezeigt werden, dass diese mehr Dokumente auf dem Desktop anhäufen, um die Sorge, wichtige Aufgaben zu vergessen, niedrig zu halten (Massey et al., 2014, S.3703). Ihrer Ansicht nach sollten Persönlichkeitsmerkmale auch bei der Entwicklung von PIM-Werkzeugen durch Personalisierungsmöglichkeiten berücksichtigt werden. Dies ist eine Position, die bereits von Boardman und Sasse eingenommen wird (Boardman & Sasse, 2004, S.589).

Jones fasst PIM-Untersuchungen in unterschiedlichen Alltagssituationen bzw. von verschiedenen Personengruppen zusammen (Jones, 2015, S.18—41). Er überblickt PIM-Arbeiten in den Kontexten: zu Hause (*PIM@Home*), in der Schule (*PIM@School*), in der Arbeit (*PIM@Work*), in der Freizeit (*PIM@Play*) und im Allgemeinen (*PIM@Large*). Studien im Kontext *PIM@Home* untersuchen z. B. die Zusammenstellung, Verwaltung und Funktion von analogen und digitalen Fotoalben bzw. Fotokollagen (Whittaker et al., 2010; Petrelli et al., 2014). Auch Studien zu Notizen an Kühlschränken, die an zukünftige Ereignisse erinnern sollen, oder es erlauben in Erinnerung an alte Ereignisse zu schwelgen, werden von Jones in diesem Kontext angeführt (L. Swan & Taylor,



2005). Jones argumentiert, dass persönliche Informationen zu Hause eben nicht nur in digitaler Form auf Laptops oder Smartphones vorliegen, sondern auch im physischen Raum vorhanden sind. *PIM@School* verdeutlicht, dass die PIM-Forschung am Verhalten von Schülern, Eltern und Lehrern interessiert ist. Diekema und Olsen untersuchen das PIM-Verhalten von Grundschullehrern (Diekema & Olsen, 2014). Finneran erforscht Faktoren, die Meta-Aktivitäten, wie das Aufbewahren oder das Zurücklassen von Information, bei College-Studenten beeinflussen (Finneran, 2008).

#### **Dimension: Methodik**

Um PIM-Verhalten empirisch zu messen und dabei eine größtmögliche ökologische Validität der Ergebnisse — Übertragbarkeit der Ergebnisse auf das Alltagsgeschehen — zu erreichen, werden auch methodisch neue Wege beschritten. PIM-Nutzerexperimente oder kontrollierte Nutzerstudien sind aufgrund der starken individuellen Voraussetzungen schwierig zu entwerfen und durchzuführen (Boardman & Sasse, 2004; Capra & Perez-Quinones, 2005; Elsweiler & Ruthven, 2007). Sie sollten mit eigenen PIM-Kollektionen realisiert werden, wobei der Anreiz des Wiederfindens intrinsisch motiviert sein und nicht von Außen vorgegeben werden sollte. Organisationsverhalten wird häufig via *guided tours* untersucht, bei denen Nutzer dem Forscher einen Einblick in ihr Organisationsverhalten geben (Malone, 1983; Kwasnik, 1991; Boardman & Sasse, 2004). Studien, die das Organisieren und Wiederfinden analysieren, werden häufig auch mit artifiziellen, fiktiven Informationen durchgeführt, was jedoch selten die gewünschten Resultate ermöglicht (Bergman & Whittaker, 2016). Elsweiler und Ruthven zeigen eine Möglichkeit, wie personalisierte PIM-Aufgaben für Laborexperimente auf der Basis von Tagebuchstudien erstellt werden können (Elsweiler & Ruthven, 2007). Bergman und Whittaker entwickeln die Untersuchungstechnik *elicited personal information retrieval* (EPIR), bei der Wiederfindensaufgaben mit eigenen persönlichen Kollektionen am eigenen Rechner bearbeitet werden (Bergman & Whittaker, 2016, S.82).

#### **Dimension: Prototypen-Entwicklung**

Im Laufe der Jahre sind zahlreiche Systeme vorgestellt worden, in deren Entwicklung die Erkenntnisse empirischer Studien eingeflossen sind und das Ziel haben, das PIM-Verhalten der Nutzer besser zu unterstützen (Dumais et al., 2003; Jones, Munat et al., 2005; Cutrell et al., 2006; Chau et al., 2008). Eines der bekanntesten Projekte stellt sicherlich das *Stuff I've Seen*<sup>4</sup> (SIS) Projekt von Dumais et al. dar (Dumais et al., 2003). SIS ist eine Desktopsuchmaschine, die das Suchen und Verwalten von verschiedenen Dateitypen (E-Mail, Web-Bookmarks, Dateien etc.) erlaubt. In einer großangelegten Logstudie zur Evaluation der Systemnutzung, können sie ebenfalls interessante Erkenntnisse in Bezug auf die Nutzung der Schlüsselwort-Suche beim Wiederfinden sammeln. Sie zeigen, dass Nutzer die Suche hauptsächlich (in 74% der Fälle) zum Wiederfinden von E-Mails nutzen. In Bezug auf das Alter der gesuchten Nachrichten waren 46% der Nachrichten jünger als einen Monat, folglich waren 54% der gesuchten Nachrichten älter als einen Monat (Dumais et al., 2003). Während Jones dies als Indiz dafür sieht, dass unsere Kollektionen altern und Information wiedergefunden werden, die sehr alt sein können (Jones, 2013, S.29), interpretieren Bergman und Whittaker diesen Umstand als einen Fokus auf Neuheit (Bergman & Whittaker, 2016, S.86). Die Suchanfragen der Nutzer enthalten zu 25% der Fälle den Namen des Senders, was ein gutes Erinnerungsvermögen an den Sender voraussetzt (Dumais et al., 2003). Dies steht im Gegensatz zur Erkenntnis von Elsweiler, Baillie und Ruthven, welche in ihrer Studie nur ein schlecht ausgeprägtes Erinnerungsvermögen an Sendernamen feststellen können (Elsweiler et al., 2008, S.17). Bergman und Whittaker versuchen diese eher konträren Ergebnisse mit der Verschiedenheit der Methodik zu erklären, die im einen Fall naturalistisch im anderen Fall experimentell war (Bergman & Whittaker, 2016, S.87). Ein weiteres System stellt das von Cutrell und Kollegen entwickelte System *Phlat* dar, welches dem Nutzer eine Kombination aus Suche, Filtermöglichkeiten, Browsingstrukturen und Tagging bietet. Auch in der Studie dieser Autoren stellen Personennamen mit

<sup>4</sup>Die Publikation hierzu besitzt über 900 Zitationen auf Google Scholar und hat 2015 den *SIGIR Test of Time Award* verliehen bekommen, welcher für Veröffentlichungen mit hohem Einfluss verliehen wird.

etwa 18% die meisten Suchanfragen (Cutrell et al., 2006, S.267). Zur Unterstützung der Navigation in Ordnerstrukturen entwickeln Chau, Myers und Faulring *Feldspar* (Chau et al., 2008). Eine Weiterentwicklung von Dateisystemen zur verbesserten Navigation in diesen wird von Fitchett, Cockburn und Gutwin vorgenommen (Fitchett et al., 2014). Neben konkreten Werkzeugen werden auch Designansätze für die Entwicklung von PIM-Werkzeugen konzipiert. Bergman, Beyth-Marom und Nachmias entwickeln den *User-subjective approach to personal information management systems design*, der dem Umstand Rechnung tragen soll, dass PIM-Verhalten äußerst individuell ist und daher einer erhöhten Personalisierung bedarf (Bergman, Beyth-Marom & Nachmias, 2008).

#### **Dimension: Untersuchungsgegenstand**

Aus den letzten Abschnitten ist deutlich geworden, dass es hinsichtlich des Untersuchungsgegenstands wenig Grenzen in der PIM-Forschung gibt. Angefangen von analogen Objekten wie Papierdokumenten (Malone, 1983; Kwasnik, 1991; Marshall & Bly, 2005), Fotos (Petrelli et al., 2014) oder Notizen an Kühlschränken (L. Swan & Taylor, 2005), über digitale Informationsobjekte, wie Digitalfotos (Whittaker et al., 2010), Webseiten (Jones et al., 2001; Cockburn & McKenzie, 2001; Bruce et al., 2004b) und E-Mails (Elsweiler et al., 2008; Whittaker et al., 2011) bis hin zu Einträgen auf Social-Media-Plattformen (Zhao et al., 2013; Marshall & Shipman, 2014) (Näheres hierzu in Abschnitt 3.2). All diese Informationstypen bzw. Kollektionen werden aus der Perspektive des PIM untersucht. Jedoch stellt Jones zu Recht fest: „There is a natural tendency to focus on one form of information in order to manage the scope of inquiry“ (Jones, 2007, S.46). Die einzelnen Informationstypen werden oft in Isolation als einzelne Informationskollektionen betrachtet, der gesamte *personal space of information* einer Person wird selten untersucht. Es existieren jedoch Studien, die das PIM-Verhalten mit mehreren Informationstypen gleichzeitig analysieren (Boardman & Sasse, 2004; Ravasio et al., 2004; Jones, Phuwanturak et al., 2005; Jones et al., 2014). Eine der bekanntesten Cross-Informationstyp-Studien stellt sicherlich die Arbeit von Boardman und Sasse mit dem Titel *Stuff Goes into the Computer and Doesn't Come out* dar (Boardman & Sasse, 2004). In einer aus zwei Phasen bestehenden Studie, analysieren Boardman und Sasse das PIM-Verhalten in Bezug auf verschiedene Informationstypen und über einen längeren Zeitraum hinweg. Sie verwenden semi-strukturierte Interviews und ein Langzeit-Tracking der Ordnerstrukturen in Verbindung mit einer Tagebuchstudie (Boardman & Sasse, 2004, S.583). Sie stellen fest, dass PIM-Strategien selbst bei den gleichen Personen stark in Abhängigkeit vom Informationstyp schwanken können (Boardman & Sasse, 2004, S.589). Weitere Erkenntnisse sind (Boardman & Sasse, 2004, S.590):

- Ob und wie stark sich Nutzer mit PIM-Strategien befassen, hängt vom wahrgenommenen Wert der Information ab. Je mehr Ressourcen des Nutzers in die Erstellung eines Dokumentes fließen, desto eher ist er dazu bereit, dieses auch aufzubewahren und zu organisieren.
- Der Aufwand für die Organisation von Informationsobjekten hängt von der Wahrscheinlichkeit und Art der erneuten Verwertung/der Wiederfindensstrategie ab.
- Nutzer äußern eine Präferenz für navigationsbasiertes Wiederfinden in Ordnerstrukturen. Diese Strukturen ändern sich selten, um Vertrauen in diese Strukturen durch ausreichend bestehende kontextuelle Information aufbauen zu können.

Nachdem gezeigt wurde, womit sich PIM-Forschung beschäftigt, respektive welche Dimensionen sie besitzt, werden die PIM-Praktiken im Detail vorgestellt.

### **2.2.2. PIM-Praktik Aufbewahren**

Aufbewahrensverhalten ist von starkem Individualismus geprägt und reicht von einem Extrem des Spektrums, dem völligen Verzicht auf Aufbewahren, bis hin zum anderen Extrem, der Aufbewahrung jeglicher Information, bei dem versucht wird, Informationen in eigene Organisationsstruktu-

ren zu überführen (Jones, 2007, S.37). Wenn Nutzer relevante Information finden, oder sie dieser zufällig begegnen, beginnt der bereits angesprochene Dialog mit dem zukünftigen Selbst. Diesem Selbst stellen wir Fragen, die wie folgt formuliert sein können:

- Wird diese Information relevant sein und sollte ich sie deshalb aufbewahren?
- In welchem Format und auf welche Art und Weise sollte ich die Information aufbewahren?
- Kann ich die Information, auch ohne aktiv tätig zu werden, in Zukunft einfach wieder aufrufen?
- Was mache ich mit Information, die ich vielleicht noch näher begutachten muss, um diese Entscheidung treffen zu können, mir aber in der aktuellen Situation die Zeit dazu fehlt?
- Wie erinnere ich mich daran, relevante Information gesehen zu haben, um anschließend zu ihr zurückzukehren oder sie zu prüfen?

Entscheidungen und Handlungen, die mit dem Abgleich (*mapping*) von gerade gesuchter oder zufällig gefundener Information und zukünftigem Informationsbedürfnis in Beziehung stehen, werden als Aktivitäten des Aufbewahrens (Keeping-Aktivitäten) bezeichnet (Jones, 2007, S.35). Auch negativ ausfallende Relevanzentscheidungen, also das Ignorieren der Information, oder die Auffassung, dass man leicht zu einer Information zurückkehren kann ohne eine bestimmte Handlung vornehmen zu müssen, gehören zu den Keeping-Aktivitäten.

Aufbewahrensentscheidungen sind sehr komplex und fehleranfällig, da sie von uns eine Vorhersage für die Zukunft erfordern. Die von Bruce geprägte PAIN-Hypothese beschreibt genau diese Schwierigkeit, Informationsbedürfnisse zu antizipieren (Bruce, 2005). Sie wurde bereits in Abschnitt 2.1.2 angesprochen. Aufbewahrensentscheidungen sind aber nicht nur deswegen so komplex, weil sie von uns einen Blick in die Zukunft verlangen, sondern auch deswegen, weil Aufbewahren bzw. das Ignorieren von Information mit Kosten verbunden ist (Jones, 2004, 2007; Bergman & Whittaker, 2016). Die Idee, dass Aufbewahrensverhalten mit Kosten einhergeht, entspringt der Modellierung des Keeping-Verhaltens auf der Basis der *theory of signal detectability* (TSD) (Jones, 2004). Jones beschreibt falsche Aufbewahrensentscheidungen mit Begrifflichkeiten aus dem Information Retrieval als *false positives* — wenn man nicht mehr benötigte bzw. nicht relevante Information aufbewahrt — oder *false negative* — wenn man eigentlich relevante Information nicht aufbewahrt (*miss*).

Kosten entstehen durch den Trade-off zwischen zu viel Aufbewahren (*overkeeping*) und zu wenig Aufbewahren (*underkeeping*). *Overkeeping* und *underkeeping* können sich vor allem in zwei Situationen als Problem erweisen: bei der Organisation und bei der Wiederverwertung der Information (Bergman & Whittaker, 2016, S.49). Organisationskosten entstehen durch den Aufwand, der geleistet werden muss, um Information beim Aufbewahren in bestehende Organisationsstrukturen einzugliedern bzw. diese zu erstellen. Ein Prozess der, folgt man der Literatur, fehleranfällig und aufwändig ist (Lansdale, 1988; Malone, 1983; Whittaker & Sidner, 1996). Er ermöglicht es uns aber, die richtige Information zum richtigen Zeitpunkt zur Verfügung zu haben. Kosten entstehen dann, wenn sehr viel Information aufbewahrt wird, da man beim Wiederfinden sehr viele Informationen begutachten muss, um zur gesuchten Information zurückkehren zu können. Alte, eigentlich nicht mehr relevante Information, kann die Aufmerksamkeit des Nutzers in Anspruch nehmen und vom eigentlichen Wiederfindensziel ablenken (Jones, 2004; Whittaker et al., 2010). Die Kosten, Information nicht aufzubewahren, können jedoch noch höher sein, da der Prozess zur Rückkehr aufwändiger wird und den Nutzer möglicherweise erfolglos und frustriert zurücklässt. Studien im Webkontext zeigen, dass Nutzer nicht dazu in der Lage sind, zu Information zurückzukehren, die sie in früheren Sitzungen gefunden haben (Wen, 2003). Beim Aufbewahren versucht der Nutzer also stets die drei Faktoren: (i) die Kosten des Aufbewahrens, (ii) den wahrgenommenen Wert der Information und (iii) die Kosten nicht verfügbarer Information in eine Balance zu bringen (Bruce,

2005).

Wen stellt fest, dass Nutzer, die es versäumen Information aufzubewahren, sich aber in Situationen des Informationsbedarfs daran erinnern relevante Information zu diesem Bedarf gesehen zu haben, besonders frustriert sind, wenn sie diese nicht erneut lokalisieren können (Wen, 2003). Er bezeichnet die Situation der nachträglich erkannten Relevanz von Information als *post-valued recall* oder *post retrieval value* (Wen, 2003).

Wie können *overkeeping* und *underkeeping* in eine Balance gebracht werden? Diese Frage ist bei der zunehmenden Flut an Informationen, sowie der Fragmentierung der eigenen Dateien über mehrere Endgeräte und Cloud-Dienste hinweg immer schwieriger zu beantworten. Folgt man Bergman und Whittaker tritt der *information overload* dann auf, wenn Nutzer in ihrer gegenwärtigen Situation nicht genügend Zeit haben, die Bedeutung von Information einzuschätzen (Finneran, 2008; Grevet et al., 2014; Bergman & Whittaker, 2016). Nutzer beginnen noch nicht adäquat verarbeitete respektive beurteilte Informationen anzuhäufen, um diese einer zeitversetzten Bewertung (auch *deferred evaluation* genannt) unterziehen zu können. Die Orte/Archive, an denen diese Information abgespeichert sind, werden jedoch selten besucht — Tauscher und Greenberg zeigen z.B., dass weniger als 50% der gesetzten Bookmarks genutzt werden (Tauscher & Greenberg, 1997) —, da häufig die nötige Erinnerungsfunktion, wie sie z. B. von E-Mails im Posteingang oder Dateien auf dem Desktop geleistet wird, nicht vorhanden ist (Barreau & Nardi, 1995) (Bergman & Whittaker, 2016, S.63). Daraus resultiert, dass Nutzer vergessen diese Information gefunden zu haben und somit ebenso vergessen, diese auf Relevanz zu prüfen. Auch bei bereits als relevant identifizierter Information kann es sein, dass diese abgelegt und vergessen wird. Jones und Kollegen bezeichnen dies als Out-of-sight-out-of-mind-Problem (Jones, Phuwanartnurak et al., 2005). Das Problem der verschobenen oder nachträglichen Evaluation von Information tritt auch dann auf, wenn Nutzer in ihrer Informationssuchaufgabe unterbrochen werden. Sie nutzen die aufbewahrte Information dann, um den Prozess weiterzuführen. Nicht verarbeitete Information wird also immer mehr angehäuft, was ebenfalls zum Phänomen des *overkeeping* führt. Eine Lösung für dieses Problem können Erinnerungsfunktionen zum Besuch der Archive darstellen.

Die Forschung hat in den letzten Jahrzehnten auf der Basis vieler empirischer Untersuchungen mit unterschiedlichen Informationstypen zahlreiche Faktoren aufgedeckt, die unsere Aufbewahrenscheidungen und -verhaltensweisen beeinflussen. Im Folgenden wird eine Reihe an Faktoren aufgezählt, die aus mehreren Forschungsarbeiten zusammengetragen sind:

- Informationstyp bzw. Informationskanal z. B. E-Mail vs. Webseite (Boardman & Sasse, 2004)
- Eigenschaften der Information/Informationsebenen (informativ vs. handlungsbedürftig, Einzigartigkeit etc.) (Barreau & Nardi, 1995; Elsweiler et al., 2008; Whittaker et al., 2010; Bergman & Whittaker, 2016)
- Metadaten (z. B. Ersteller, Titel, Alter) der Information (Kwasnik, 1991; Elsweiler et al., 2008)
- Wahrgenommener Wert der Information (Boardman & Sasse, 2004)
- Zukünftige Funktion bzw. zukünftiger Kontext (Barreau & Nardi, 1995; Jones et al., 2001)
- Persönlichkeitsmerkmale des Aufbewahrenden (Boardman & Sasse, 2004; Massey et al., 2014)
- Auf der Basis der First-Impression-Hypothese: Art, wie Information initial gefunden wurde (bewusst vs. serendipityartig) (Jones, 2013; Elsweiler et al., 2008)
- Art, wie Information organisiert wird (z. B. Existenz bestehender Strukturen) (Boardman & Sasse, 2004)
- Wahrgenommene Einfachheit, mit der man zur Information zurückkehren kann (Bruce et al.,

2004b)

Die Liste führt vor allem Faktoren an, die Aufbewahren rational motiviert sehen. Aufbewahren muss jedoch nicht immer zweckgebunden bzw. rational sein. Auch rein sentimentale Beweggründe motivieren Menschen dazu, Dinge aufzubewahren. Marshall und Bly untersuchen das Aufbewahrensverhalten von Probanden mit Ausschnitten aus digitalen und analogen Zeitungen und stellen fest, dass 10% der Ausschnitte einfach dazu dienen, Erinnerungen zu wecken, oder eine Sammlung an schönen Artikeln zu erstellen (Marshall & Bly, 2005, S.116). Aufbewahren kann als mentale Stütze erfolgen (Erinnerungsfunktion), eine Art der Selbstpräsentation sein oder ohne jeglichen plausiblen Grund geschehen (Jones, 2007, S.38). Faktoren, die das Aufbewahren beeinflussen, überschneiden sich natürlich mit Faktoren, die Wiederfinden charakterisieren, da beides komplementäre Aktivitäten sind (Teevan, Capra & Perez-Quinones, 2007, S.24). In der Praxis kann Aufbewahren je nach Informationstyp sehr unterschiedliche Formen annehmen. Es reicht von der Nutzung der Ordnerstruktur eines Dateisystems oder dem Ordnersystem eines E-Mail-Clients, dem Ablegen einer Datei auf dem Desktop, dem sich Zusenden von wichtigen Webseiten per E-Mail, dem Ausdrucken von Webseiten, oder auch der Nutzung von Tagging-Funktionalität, wenn Systeme dies erlauben.

Schließlich stellt sich die Frage nach Lösungsmöglichkeiten, um das Problem der Aufbewahrensentscheidung einfacher zu gestalten. Ansätze wie *keep everything* scheinen aufgrund der hohen Organisations- und Retrievalkosten ebenso wenig erfolgversprechend, wie der Ansatz *keep nothing (and search everything)* — wie er z. B. von Russell (Russell & Lawrence, 2007) vertreten wird — der ebenfalls Schwächen aufweist. Alternativen stellen der Keep-Automatically-Ansatz dar, der z.B. eine Clusteranalyse der Informationsobjekte vornimmt und basierend auf deren inhaltlicher Ähnlichkeit ein Aufbewahren und Kategorisieren in bereits vorhandenen Ordnern durchführt (Jones, 2007, S.53). Mit Verfahren des automatischen Aufbewahrens wurde bereits bei Systemen wie dem E-Mail-Client *MailCat* experimentiert (Segal & Kephart, 1999). Auch neuere Ansätze wie der *PiMxT*-Prototyp, entwickelt von Abela und Staff, arbeiten mit automatischen Keeping-Verfahren auf der Basis von Clustering-Verfahren (Abela & Staff, 2016). Grbovic et al. untersuchen ein automatisches Klassifizieren von E-Mails einer Web-Mail-Anwendung auf der Basis von latenten Kategorien, wobei sie feststellen, dass sechs Kategorien ausreichen, um die meisten Mails aller Nutzer zu klassifizieren (Grbovic et al., 2014). Das automatische Kategorisieren von E-Mails ist jedoch mit Problemen verbunden, da die Gefahr besteht, dass Mails nie in die Aufmerksamkeitsebene des Nutzers gelangen und somit einfach weg-archiviert werden. Weitere Ansätze versuchen eine Verbesserung des Ordneransatzes mithilfe von IR-Verfahren, wie dies z. B. auch bei Googles E-Mail-Client *GMail*<sup>5</sup> realisiert ist. In *GMail* existieren Ordner nicht physisch sondern lediglich als Tags, die für eine Filtersuchanfragen dienen (Grbovic et al., 2014). Eine weitere Lösung könnte der Ansatz *keep smarter* darstellen, der auf der Grundlage von Informationsfiltermethoden aufbewahrenswerte Informationsobjekte vorschlägt und schließlich den Nutzer beim Aufbewahrensprozess mit einbindet (Jones, 2007, S.45).

### 2.2.3. PIM-Praktiken Organisieren und Verwalten

Aus der PIM-Perspektive bietet Twitter wenig Optionen, um Tweets zu verwalten. Lediglich das Löschen von eigenen gesendeten Nachrichten bzw. das Entfavorisieren — das Löschen aus der Favoritenliste — können als Verwaltungshandlungen interpretiert werden. Organisationsstrukturen existieren in diesem Sinne folglich nicht, wobei man aber die eigenen gesendeten Tweets und die Favoritenliste als *piles*, also Haufen, interpretieren kann (Malone, 1983).

Organisationsverhalten (*organizing*) wird von Jones wie folgt definiert:

<sup>5</sup><https://www.google.com/gmail/>

„Decision-making and actions relating to the selection and implementation of a scheme of organization and representation for a collection of information items. Decisions can include: (1) How should items in this collection be named? (2) What set of properties make sense for and help to distinguish the items in this collection? (3) How should items within this collection be grouped? Into piles or folders“ (Jones, 2007, S.39)

Während sich das Aufbewahren immer auf ein spezifisches Informationsobjekt bezieht, betrifft die Organisation oder auch die Verwaltung (*maintaining*), immer die gesamte Informationskollektion, z. B. die Kollektion aller E-Mails. Organisieren und Verwalten tritt oft in Kombination oder abwechselnd auf, jedoch besteht auch hier ein Unterschied. Während Organisationsverhalten als Ziel das Erstellen und das Einordnen von Informationsobjekten in eine Organisationsstruktur besitzt, versteht man unter *maintaining* das Löschen oder Verändern dieser Strukturen bzw. eine langfristige Pflege, wozu auch regelmäßige Backups gehören (Jones, 2007, S.40). Aufbewahren ist also eine Aktivität, die täglich auftritt. Organisieren oder Verwalten von Kollektionen hingegen gehört zur Kategorie der Meta-Verhaltensweisen, die weniger häufig auftreten. Oh fasst die wesentlichen Perspektiven auf Organisationsverhalten in vier Punkten zusammen, wobei jeweils eine exemplarische Arbeit, die die jeweilige Perspektive untersucht, genannt wird (K. E. Oh, 2013, S.66):

- Von Nutzern erstellte Organisationsstrukturen und Klassifikationssysteme, sowie Quantifizierung der Dateien, Ordner und Tiefe/Breite dieser Strukturen (Bergman et al., 2010)
- Organisation unterschiedlicher Informationstypen (Barreau & Nardi, 1995)
- Individuelle Unterschiede in persönlichen Informationsorganisationsstrategien (Boardman & Sasse, 2004).
- Kriterien/Faktoren, die das Organisationsverhalten beeinflussen (Kwasnik, 1989; Barreau & Nardi, 1995)

Anhand der PIM-Perspektive *organizing* werden die zwei Denkschulen bzw. Lager innerhalb der PIM-Community abermals deutlich. Die Vertreter des Search-everything-Lagers argumentieren, dass Ordnerstrukturen und Verwaltung von Dateien nicht länger notwendig sind (Russell & Lawrence, 2007). Der Fortschritt, den IR-Techniken erfahren haben, sowie unendliche Speichermöglichkeiten befreien uns von dem Zwang, Informationen aufzubewahren und zu organisieren. Konkret heißt es: „Folders in particular, as an organizing construct, are targeted for obsolescence“ (Jones, 2007, S.47). Jedoch gibt es eine Reihe an Argumenten seitens der Befürworter von Organisationsstruktur, die dafür sprechen, dass die Organisation von Information mehr ist, als nur den Zugang zu ihnen zu gewährleisten. Im vorherigen Abschnitt wurden bereits Faktoren genannt, wie Kontextinformation oder Erinnerungsfunktion, die durch Aufbewahren und Organisationsstrukturen geleistet werden. Nach Jones helfen sie nicht nur beim Wiederfinden (vgl. nächster Abschnitt) sondern auch dabei, neuer Information Sinn zu verleihen. Sie unterstützen somit den Prozess des *sense making* (Jones, 2007, S.45). Vorhandene Organisationsstrukturen ermöglichen es, die Bedeutung von Information, der man neu begegnet, besser beurteilen zu können. Hat man einmal die Bedeutung von Information erfasst, hilft dies auch bei ihrer Kategorisierung (Jones & Ross, 2007; Jones, Phuwanartnurak et al., 2005).

Frühe Studien zum Organisationsverhalten werden bereits im analogen Kontext bei der Organisation von Papierdokumenten im Büro durchgeführt (Cole, 1982; Malone, 1983; Kwasnik, 1989). Die Pflege einer Ordnerstruktur und deren Inhalte geht oft mit einer Re-Organisation oder einem Löschen dieser Inhalte einher. Bergman und Whittaker argumentieren, dass Nutzer fortwährend kleine Veränderungen an der Struktur vornehmen (Bergman & Whittaker, 2016, S.52). Sie beziehen sich auf die Studie von Boardman und Sasse, in der die Probanden im Durchschnitt alle drei Tage einen neuen Ordner im Dateisystem angelegt haben, sowie jeden fünften Tag einen neuen Ordner im E-Mail-Programm erzeugten (Boardman & Sasse, 2004). Einen Überblick zu PIM-Studien im Kontext der Organisation von E-Mails nehmen Whittaker, Bellotti und Gwizdka vor (Whittaker et

al., 2007). Sie fassen acht Studien aus den Jahren 1996 bis 2005 zusammen und fokussieren in ihrer Darstellung auf Management von E-Mails bzw. das Organisationsverhalten im Allgemeinen. Organisationsverhalten wird oft versucht, in Form von Nutzerklassen bzw. Organisationstypen zu beschreiben. Eine der bekanntesten Einteilungen von Organisationsstrategien bzw. Nutzerklassen im Umgang mit E-Mails stammt von Whittaker und Sidner (Whittaker & Sidner, 1996). Sie identifizieren drei Kategorien:

- *No-filers*: Nutzer, die auf die Verwendung von Ordnern komplett verzichten
- *Frequent-filers*: Nutzer, die sehr viele Ordner besitzen und diese auch regelmäßig verwenden
- *Spring-cleaners*: Nutzer, die eine periodische Säuberung des Posteingangs vornehmen

Diese Einteilung wird zehn Jahre später von Fisher et al. um die Nutzerklasse der *few folder filers* ergänzt, welche Nutzer beschreibt, die zwar wenige Ordner besitzen, aber deren Inbox auch sehr klein ist (D. Fisher et al., 2006, S.312). Abrams, Baecker und Chignell nehmen eine ähnliche Kategorisierung in Bezug auf den Umgang von Nutzern mit Webbookmarks vor (Abrams et al., 1998).

Das Bewusstsein, dass Organisationskonzepte Pflege bedürfen und verwaltet werden müssen, erlangen Nutzer, so Jones, nur bei seltenen Gelegenheiten, wie dem Ende eines Projektes: „Since maintenance activities are rarely forced by the events of a typical day, these activities often occur after work is done — which is to say rarely or not at all“ (Jones, 2008, S.157). Häufig wird der Prozess auch von außen angestoßen, wie bei Boardman und Sasse beobachtet. In ihrer Studie begannen die Probanden während der *guided tours* motiviert durch die Anwesenheit der Experimentleiter mit sogenanntem *ad hoc tidying* (Boardman & Sasse, 2004, S.588). Auch in den KFTF-Studien begannen Versuchsteilnehmer im Laufe der Studie sich selbst als chaotisch oder als digitale Messies zu bezeichnen und begannen mit Verwaltungsarbeiten (Jones, 2008, S.159).

Ein großer Bereich, der mit der Pflege oder Verwaltung von Informationskollektionen einhergeht, betrifft das Löschen von Informationsobjekten. Löschen ist etwas, das Nutzer ungern vornehmen. Viele leiden unter der von Kahneman geprägten Angst der *loss aversion*, also dem Zwang Information zu behalten, da man sich — oft aufgrund emotionaler Bindung — nicht von ihr trennen kann (Bergman & Whittaker, 2016, S.37). In Zusammenhang mit dem Löschen von Informationsobjekten aus Informationskollektionen formulieren Bergman, Beyth-Marom und Nachmias das Phänomen des *deletion paradox* (Bergman et al., 2003). Das *deletion paradox* beschreibt den Umstand, dass die Evaluation von Informationsobjekten, die eigentlich keinen Wert mehr besitzen, bei einem Löschvorgang viel Zeit und Aufwand in Anspruch nimmt, wenn sichergestellt wird, dass diese tatsächlich nicht mehr relevant sind und folglich entfernt werden können (Bergman et al., 2003, S.876). Jones beschreibt den Grund, warum man derartige Informationsobjekte überhaupt aufbewahrt — und hier bewegt man sich abermals im Bereich der Aufbewahrensentscheidung aus Abschnitt 2.2.2 — als *old magazine effect*, da der potentielle Nutzen, den die Information zu dem Zeitpunkt der Entscheidung besitzt, sichtbarer ist, als die laufenden Kosten, die durch deren Verwaltung entstehen (Jones, 2008, S.158). Der Grund, weshalb derartige Säuberungen vorgenommen werden, ist nicht ein Problem des geringen Speicherplatzes. Informationskollektionen müssen re-organisiert oder gelöscht werden, da Nutzer ein beschränktes Aufmerksamkeitsvermögen bezüglich des Raums oder der Fläche besitzen, die beim Wiederfinden verarbeitet werden kann: „For these surfaces to retain their attention-getting power, they need to be cleared of the deritus of old stuff that builds up“ (Jones, 2008, S.161). Je mehr Aufmerksamkeit also einzelnen Informationsobjekten zukommen kann, desto einfacher und schneller können Informationen wiedergefunden werden.

Im Folgenden wird das Wiederfindensverhalten näher charakterisiert.

### 2.2.4. PIM-Praktik Wiederfinden

Der Schwerpunkt der Arbeit liegt auf der PIM-Praktik des Wiederfindens, weshalb diese im Folgenden ausführlicher beschrieben wird. Neben dem tatsächlichen Akt des Wiederfindens werden auch verwandte, oft auch synonym gebrauchte Verhaltensweisen, wie das Wiederbesuchen (*Re-visitation*) oder das erneute Entdecken (*Re-discovery*) angesprochen. Diese Unterscheidung wurde bereits in der Einleitung thematisiert. *Re-visits* unterscheiden sich von Wiederfinden dahingehend, dass das erneute Besuchen kein Wiederfinden voraussetzt. Dies ist z. B. bei einer Website der Fall, die noch im Browser-Tab geladen ist. Wiederbesuche sind also zeitlich gesehen sehr nah am initialen Aufruf des Informationsobjekts. Das erneute Entdecken tritt hingegen dann auf, wenn der initiale Aufruf bereits längere Zeit zurückliegt und der Nutzer eher zufällig auf das Informationsobjekt trifft, z. B. beim Durchstöbern seiner Bookmarks. Ein besonderer Fokus wird auf das Wiederfindensverhalten in unterschiedlichen Kontexten (E-Mail, Web) liegen, wobei Studien vorgestellt werden, die dieses Verhalten via Logdatenauswertungen untersuchen.

In der Einleitung (vgl. Abschnitt 1.2) wurde Wiederfinden bereits als Prozess bzw. Verhalten definiert, bei dem man zur Information zurückkehrt, die man bereits wahrgenommen hat. Finden und Wiederfinden treten bei der Erledigung von Aufgaben oft gemeinsam auf, jedoch unterscheiden sich beide Verhaltensweisen voneinander (Teevan, Capra & Perez-Quinones, 2007, S.24). Capra, Pinney und Perez-Quinones formulieren diesen Sachverhalt als „Re-finding Is Not Finding Again“ und weisen damit darauf hin, dass Wiederfinden mehr als nur ein Nachvollziehen der Schritte des initialen Findens ist (Capra et al., 2005). Re-finding unterscheidet sich vor allem deshalb signifikant vom initialen Finden, da der Nutzer zum Wiederfindenszeitpunkt über mehr Meta- oder Kontextwissen verfügt (Teevan, Capra & Perez-Quinones, 2007, S.29). Häufig werden dabei subjektiv erstellte Organisationsstrukturen ausgenutzt (Bergman & Whittaker, 2016, S.87). Zahlreiche Studien können zeigen, dass das Wissen über das Vorhandensein einer Information zu einem weitaus frustrierenderen Nutzungserlebnis beim Wiederfinden führen kann (Wen, 2003; Teevan, Capra & Perez-Quinones, 2007; Whittaker et al., 2010).

In empirischen Studien haben sich zwei Hauptkategorien von Wiederfindensstrategien herausgebildet: Wiederfinden via manueller Navigation/Browsing durch Folgen von Hyperlinks, was auch als *orienteeing* (Teevan et al., 2004) oder *location-based finding* (Barreau & Nardi, 1995) bezeichnet wird und Wiederfinden via Nutzung von Schlüsselwort-Suche, was auch als *teleporting* (Teevan et al., 2004) oder *logical finding* (Barreau & Nardi, 1995) bezeichnet wird. In der Forschung existieren dementsprechend auch zwei Lager, die jeweils eine Wiederfindensstrategie als die bessere Vorgehensweise charakterisieren. Befürworter des Orienteering-Ansatzes argumentieren, dass dieses Verhalten von einer kognitiven Leichtigkeit geprägt ist und führen folgende Argumente an:

- Die Recognition-vs-Recall-Hypothese trifft zu, was bedeutet, dass es Nutzern leichter fällt, Eigenschaften wie Dateinamen wiederzuerkennen, als sie komplett zu erinnern.
- Der Orientierungssinn der Nutzer wird durch eine bekannte Umgebung (Ordnerstruktur/Hypertextstruktur) gefördert.
- Der Navigationskontext liefert Feedback darüber, ob der richtige Weg eingeschlagen worden ist und ob die richtige Information gefunden wurde (Jones, 2013, S.25).

*Orienteering* mag zwar langsamer sein, jedoch ist es kognitiv leichter und weniger fehleranfällig, da Wiederfinden in mehrere kleine Schritte zerlegt wird, bei denen stets *recognition* statt *recall* im Vordergrund steht. Bergman et al. argumentieren, dass die schrittweise Rückkehr zu abgelegter Information auf leichter zugänglichem, prozeduralem Wissen beruht, wohingegen für die Formulierung einer Suchanfrage schwerer zugängliches deklaratives Wissen benötigt wird (Bergman, Beyth-Marom, Nachmias, Gradovitch & Whittaker, 2008, S.20:19). Ein weiterer Vorteil des



Orienteering-Ansatzes liegt in der Erinnerungsfunktion, über die Ordner verfügen. Ordnerinhalte zeigen uns vermeintlich vergessene Inhalte und erlauben somit ein erneutes Entdecken (*Rediscovery*) von Information (Barreau & Nardi, 1995, S.41). Zusammenfassend ist navigationsbasiertes Wiederfinden einfacher, da es inkrementell abläuft und es dem Nutzer viel Kontrolle und visuelles Feedback bietet. Zahlreiche Studien konnten diese Vorteile bzw. die Präferenz seitens der Nutzer feststellen (Teevan et al., 2004; Bergman, Beyth-Marom, Nachmias, Gradovitch & Whittaker, 2008; Bergman et al., 2010). Aber auch der Navigationsansatz besitzt Nachteile. Dies ist z. B. im Kontext von E-Mail-Re-finding zu erkennen, bei dem häufig sogenannte *failed folders* auftreten, also Ordner, die zu wenige E-Mails oder zu viele E-Mails enthalten. Zusätzlich werden diese entweder nie besucht werden oder sie enthalten so viele Mails, dass die Navigation in ihnen zu komplex wird (Effekt des *overkeeping*) (Whittaker & Sidner, 1996). Im Kontext der Navigation in Dateisystemen zeigen Bergman et al., dass jeder zusätzliche Ordner, den man tiefer navigieren muss, den Suchprozess um 2 Sekunden verlängert und ab 21 Dateien pro Ordner die Suche kognitiv belastend wird (Bergman et al., 2010). Die meisten Untersuchungen, die *orienteeing* als überlegen ansehen, beziehen sich hauptsächlich auf das Wiederfinden in Dateisystemen, so dass in anderen Kontexten nicht uneingeschränkt von einer Überlegenheit des Navigationsansatzes ausgegangen werden kann.

Die Befürworter des Teleporting-Ansatzes argumentieren, dass Schlüsselwortsuche sehr schnell ist, da man sich nur an wenig Meta-Informationen des gesuchten Informationsobjekts erinnern muss, um sie in einer Suchanfrage zu verwenden und, was vielleicht noch viel wichtiger ist, es den Nutzer komplett davon befreit, Organisationsstrukturen aufzubauen und zu verwalten (Russell & Lawrence, 2007).

Wiederfinden wird von ähnlichen Faktoren beeinflusst wie das Aufbewahren (vorgestellt in Abschnitt 2.2.2). Besonders die dort erwähnte First-Impression-Hypothese wird beim Re-finding-Verhalten evident. Die Charakteristika eines Informationsobjektes, die initial dazu genutzt werden, um seine Relevanz zu beurteilen, respektive diejenigen Charakteristika, die die Aufmerksamkeit des Nutzers auf sich ziehen und vom Nutzer verwendet werden, um eine Priorisierung unter den einzelnen Informationsobjekten im Informationsraum vorzunehmen, sind auch diejenigen Charakteristika, die später gut erinnert werden und als Terme in einer Suchanfrage genutzt werden (Jones, 2013, S.28). Dies konnte sowohl in den KFTF-Studien als auch in einer Studie von Elweiler, Bailie und Ruthven festgestellt werden (Jones et al., 2001; Jones, 2008; Elweiler et al., 2008). Auch Ergebnisse der Studie von Dumais und Kollegen — in Abschnitt 2.2.1 wurde bereits beschrieben, dass in der Evaluationsstudie zum SIS-System mit 76% E-Mails diejenigen Informationstypen waren, die am häufigsten wiedergefunden wurde, wobei das Wiederfinden hauptsächlich durch Schlüsselwort-Suche gelöst wurde und die meisten Terme Sendernamen enthielten — bestätigen diesen Eindruck (Dumais et al., 2003). Die First-Impression-Hypothese ist eine Mischung aus Faktoren des initialen Auffindens sowie den Eigenschaften des Informationsobjekts, die das Wiederfinden beeinflussen.

Im Folgenden wird Wiederfinden im Umgang mit E-Mails und im Webkontext näher betrachtet, da beide Informationsarten eine Nähe zu Tweets in Bezug auf deren Informationscharakteristika besitzen.

#### 2.2.4.1. Wiederfinden von E-Mails

In den vorherigen Abschnitten wurden bereits einige Charakteristika des Wiederfindens von E-Mails genannt. In Abschnitt 2.1.3 wurde auch auf die Besonderheit von E-Mails als Informationstyp hingewiesen, da sie sowohl informativ sein können als auch Handlungsbedarf erfordern können, was sich auf den Umgang mit ihnen beim Aufbewahren und Wiederfinden auswirkt. Das Wieder-

finden von E-Mails ist auch deshalb speziell, da E-Mail-Anwendungen häufig sowohl das Erstellen von Ordnerstrukturen erlauben als auch einen eigenen Index besitzen. Dies ermöglicht sowohl das Wiederfinden via *orientteering* als auch via *teleporting*. In Abschnitt 2.2.3 wurde bereits die Überblicksarbeit von Whittaker, Bellotti und Gwizdka erwähnt (Whittaker et al., 2007). Diese fokussiert jedoch sehr stark auf Aufbewahrens- und Managementverhalten der Nutzer, so dass keine Aussagen über deren Wiederfindensverhalten getätigt werden. Wenige Arbeiten setzen sich explizit mit dem Wiederfinden von E-Mails auseinander.

In der bereits schon mehrfach erwähnten Studie von Elsweiler, Baillie und Ruthven untersuchen diese in einer kontrollierten Laborstudie mit 34 Probanden deren Erinnerungsvermögen bei der Durchführung von Wiederfindensaufgaben in ihren eigenen E-Mail-Kollektionen (Elsweiler et al., 2008). Dabei variieren sie die Art der Aufgabe (*lookup*, *item*, *multi-item*) sowie das Alter der gesuchten E-Mails (Elsweiler et al., 2008, S.21:11). Generell können sich Nutzer sehr gut an ihre E-Mails erinnern, wobei E-Mail-Inhalt sowie Sendegrund am häufigsten erinnert werden (Elsweiler et al., 2008, S.21:30). Ferner stellen sie fest, dass das Wiederfindensszenario die erinnerten Charakteristika der gesuchten E-Mail stark beeinflusst (Elsweiler et al., 2008, S.21:30). Dieser Zusammenhang wurde weiter oben schon ähnlich diskutiert: Eigenschaften des Informationsobjekts, an die man sich erinnert, beeinflussen die Wiederfindensstrategie.

Wiederum Elsweiler, Baillie und Ruthven untersuchen, welche Faktoren das Wiederfinden von E-Mails zu einer schwierigen, zeitaufwändigen Aufgabe machen (Elsweiler, Baillie & Ruthven, 2011). Konkret beziehen sie sich dabei auf Faktoren aus drei Ebenen: die Ebene des Aufgabenkontextes (Häufigkeit, mit der man die Aufgabe durchführt, Dauer seit der man die gesuchte Information gesehen hat etc.), die Gruppenebene (Aufbewahrens- und Organisationsverhalten usw.) sowie die individuelle Ebene (z. B. Anzahl der E-Mails in der Inbox etc.). Sie berechnen ein logistisches Regressionsmodell, das wesentliche Features, die zur Schwierigkeit beim Wiederfinden beitragen, offen legt (Elsweiler, Baillie & Ruthven, 2011, S.573). Drei der fünf Variablen beziehen sich auf den Aufgabenkontext und stehen in besonderem Zusammenhang mit der Zeit, die seit dem Erhalt und dem letzten Öffnen der E-Mail vergangen ist (Elsweiler, Baillie & Ruthven, 2011, S.574). Kontraintuitiv scheint die Tatsache zu sein, dass ein Erinnern mehrerer Empfänger dazu führt, dass eine Wiederfindensaufgabe als schwieriger wahrgenommen wird, als eine Aufgabe, bei der man nur einen Empfänger erinnert (Elsweiler, Baillie & Ruthven, 2011, S.574).

Elsweiler, Harvey und Hacker (Elsweiler, Harvey & Hacker, 2011) sowie Whittaker et al. (Whittaker et al., 2011) analysieren Wiederfindensverhalten auf der Basis von lebensnahen Logstudien. Elsweiler, Harvey und Hacker fokussieren zunächst auf die methodische Herausforderung, die mit der Identifikation von Wiederfindensverhalten in Logs einhergeht, da die Nutzung eines E-Mail-Clients unterschiedliche Gründe — Aufgaben- oder Kontaktmanagement — haben kann, was die Identifikation von Wiederfindensaktionen schwieriger gestaltet. Zur Identifikation von Wiederfindensaktionen vergleichen sie zahlreiche qualitative (Kodierung) und quantitative (maschinelles Lernen) Verfahren (Elsweiler, Harvey & Hacker, 2011). Ihr bestes Verfahren identifiziert schließlich bis zu 15% der Sessions als Wiederfindenssitzungen (Elsweiler, Harvey & Hacker, 2011, S.46). Ihre Daten zeigen, dass 55% der E-Mails erneut geöffnet werden, wobei dies bei den meisten (36%) innerhalb einer Stunde der Fall ist. 7% der erneut geöffneten Nachrichten sind jedoch über ein Jahr alt. Die älteste Nachricht war sogar über 6 Jahre alt (Elsweiler, Harvey & Hacker, 2011, S.37). Es zeigt sich jedoch auch, dass erneut geöffnete Nachrichten häufiger wiedergefunden werden. Wiederfindenssitzungen dauern im Median vier Minuten, können aber über eine Stunde (71,7 Min.) in Anspruch nehmen. Elsweiler, Harvey und Hacker charakterisieren das Wiederfinden als Herausforderung, was anhand der *message uncertainty ratio*<sup>6</sup> zu erkennen ist, da bei der Suche nach der richtigen

<sup>6</sup>Message uncertainty ratio MUR=Geklickte E-Mails/Einzelne geklickte E-Mails. Erfolgen in einer Sitzung sechs Klicks auf vier unterschiedliche E-Mails, beträgt die  $MUR = \frac{6}{4} = 1,5$ .

E-Mail einzelne Nachrichten immer wieder geöffnet werden (Elsweiler, Harvey & Hacker, 2011, S.43). Folgt man den Autoren, kann in dieser Studie eine Präferenz der Wiederfindensstrategie *orientteering* ausgemacht werden, da in 72% der Fälle Ordner für das Wiederfinden genutzt werden. Sie stellen fest, dass für den Fall, dass Nutzer Ordner benutzen und pflegen, diese auch häufig zum Wiederfinden verwendet werden. Die *folder uncertainty ratio* (FUR) zeigt eine gewisse Desorientierung seitens der Nutzer, da diese mehrmals die gleichen Ordner besuchen, wenn sie einen E-Mail wiederfinden wollen. 94% der Wiederfindenssitzungen enthalten zudem auch Suchanfragen. Diese sind jedoch sehr kurz, bestehen oft nur aus einem Wort (90% nur ein einzelner Term) und werden an die Default-Suchfelder Absender oder Betreff gerichtet (Elsweiler, Harvey & Hacker, 2011, S.41-42).

In einer späteren Studie untersuchen sie die Suchanfragemuster schließlich im Detail (Harvey & Elsweiler, 2012). In Anlehnung an die Arbeit von Teevan et al. (Teevan, Adar et al., 2007) analysieren Harvey und Elsweiler den Zusammenhang zwischen der Art der Suchanfrage beim Wiederfinden (handelt es sich um eine identische oder veränderte Suchanfrage) und den erneut geklickten Links (identisch vs. verändert) (Harvey & Elsweiler, 2012, S.30). Sie stellen eine vergleichsweise hohe Anzahl (20% im Vergleich zu 4% bei der Websuche (Teevan, Adar et al., 2007)) an wiederholten, identischen Suchanfragen fest, bei denen jedoch keine Überschneidungen an Klicks auf E-Mails zu vorherigen Sitzungen besteht (Harvey & Elsweiler, 2012, S.31). Ferner stellen sie fest, dass bei den geklickten E-Mails eine hohe Senderüberlappung besteht, was bedeutet, dass Nutzer sich bei der Suche auf den Sender der E-Mail beziehen. Insgesamt ist ein starker Fokus auf Personen/Sender in ihren Suchanfragen zu erkennen (Harvey & Elsweiler, 2012, S.34).

Whittaker et al. vergleichen in ihrer Logstudie die beiden Wiederfindensstrategien *preparatory retrieval* — worunter die Nutzung von Ordnerstrukturen, also der navigationsbasierte Wiederfindensansatz, zu verstehen ist — und den *opportunistic approach* — worunter die Verhaltensweisen Sortieren, Scrollen und Suchen eingegliedert werden — bei 345 Probanden über mindestens 2 Monate. Als Untersuchungsgegenstand dient der webbasierte E-Mail-Client *Blue-Mail* (Whittaker et al., 2011). Insgesamt analysieren sie 85 000 Wiederfindensaktionen, wobei jeder Nutzer im Durchschnitt in etwa 61 Wiederfindenssituationen involviert ist, die im Durchschnitt etwa 70 Sekunden dauern (Whittaker et al., 2011, S.3454). Wiederfinden via *orientteering* findet nur in 13% der Fälle statt, wobei die Nutzergruppe der exzessiven Ordnernutzer (*high filers*) den navigationsbasierten Ansatz bevorzugt. Am häufigsten (62%) werden E-Mails wiedergefunden, indem Nutzer den Posteingang abwärts scrollen (Whittaker et al., 2011, S.3454). Oft treten jedoch die Suche, das Sortieren und schließlich das Scrollen gemeinsam auf. In Bezug auf einen Vergleich der Effizienz der beiden Retrieval-Strategien können Whittaker und Kollegen zwar zeigen, dass *high filers* weniger Schritte bis zum Wiederfinden tätigen müssen, sie stellen jedoch auch fest, dass diese Gruppe nicht schneller ist, als andere Nutzergruppen (Whittaker et al., 2011, S.3455). Eine weitere Hypothese, die besagt, dass Nutzer, die hauptsächlich den navigationsbasierten Ansatz verwenden, erfolgreicher Wiederfinden, musste ebenfalls abgelehnt werden (Whittaker et al., 2011, S.3455). Generell zeigt sich beim Wiederfinden von E-Mails eher eine Nutzerpräferenz zur Lokalisierung der gesuchten E-Mail im Posteingang via Scrolling oder der Nutzung der Schlüsselwortsuche.

Aktuelle Studien thematisieren den Search-everything-Ansatz und versuchen das Nutzungserlebnis beim Umgang mit E-Mails durch algorithmische Neuerungen bei der E-Mail-Suche zu verbessern (Carmel et al., 2015; G. Wang et al., 2016). Carmel et al. argumentieren, dass die klassische, chronologische Anzeige von E-Mail-Suchergebnissen zu hinterfragen ist und schlagen einen auf einem Learning-to-Rank-Verfahren aufbauenden Algorithmus namens *REX* vor, der E-Mail-Suchergebnisse nach Relevanz ordnet (Carmel et al., 2015, S.289).

Wie sieht das Wiederfindensverhalten der Nutzer in sich ständig verändernden, dynamischen Umgebungen wie dem Web — Twitter ist wohl noch viel dynamischer — aus? Dies wird im Folgenden

durch die Darstellung des Wiederfindensverhaltens im Web thematisiert.

#### 2.2.4.2. Wiederfinden und Wiederbesuch von Webseiten

Die ersten Studien zum Wiederfindensverhalten im Web stellen die *Keeping Found Things Found*-Studien dar (Jones, 2008). In diesen Arbeiten kann eine Präferenz von Wiederfindensstrategien festgestellt werden, die keine offenkundige Aufbewahrensaktion (*do nothing methods*) voraussetzen (Bruce et al., 2004a). In einer Studie des KFTF-Projekts von Bruce, Jones und Dumais werden in 76% der Wiederfindenssituationen Strategien benutzt, die kein Aufbewahren voraussetzten. Nutzer kehren zu Webseiten zurück, indem sie die URL der Seite direkt in den Browser eingeben (42%), erneut danach suchen (18%), oder Links von anderen Webseiten folgen (16%) (Bruce et al., 2004a). Methodisch gesehen werden in den KFTF-Studien vorwiegend Experimente oder qualitative Studien durchgeführt. Vergleichbare Studien führen Capra und Perez-Quinones durch.

Zahlreiche Studien analysieren das Wiederbesuchsverhalten von Webseiten (Tauscher & Greenberg, 1997; Cockburn & McKenzie, 2001; Obendorf et al., 2007; Adar et al., 2008). Diese Studien werden hauptsächlich als Logstudien mit adaptierten Browsern durchgeführt. Ob der Wiederbesuch (*Re-visitation*) einer Website tatsächlich als Wiederfinden charakterisiert werden kann, ist durchaus diskussionswürdig. Obendorf et al. argumentieren, dass es bei erneutem Seitenaufrufen vor allem auf die Motivation des Nutzers ankommt, ob es sich um Wiederfinden handelt oder nicht (Obendorf et al., 2007, S.601). Besucht dieser eine Seite erneut, um den gleichen Inhalt noch einmal zu lesen, kann dies als Wiederfinden interpretiert werden. Besucht der Nutzer eine Seite, um aktualisierte Inhalte zu lesen, ist dies kein Re-finding-Verhalten. Vor allem bei erneuten Besuchen, die innerhalb eines Zeitfensters von wenigen Minuten fallen und nur durch die Nutzung des Zurück-Buttons des Browsers wieder besucht werden, handelt es sich wohl nicht um Re-finding im eigentlichen Sinne. Diese Studien zeigen, dass die Häufigkeit von erneuten Seitenaufrufen mit 41% (Obendorf et al., 2007) bis zu 81% (Tauscher & Greenberg, 1997) sehr hoch ist und Nutzer Informationen weitaus häufiger wieder besuchen als man annehmen würde. Manche Seiten werden dabei auch häufiger erneut besucht. Die am häufigsten genutzten Rückkehrstrategie stellen der Zurück-Button des Browsers und das Folgen von Links dar. Erstaunlicherweise werden Bookmarks oder der Browser-Verlauf nur selten (in 0,2% der Fälle) genutzt (Obendorf et al., 2007, S.604). Zudem liegt ein starker Fokus auf Aktualität bzw. kürzlich angesehenen Webseiten. Bei Obendorf et al. finden 73% aller Wiederbesuche innerhalb einer Stunde statt, nur 7,6% treten nach einer Woche auf (Obendorf et al., 2007, S.603). Eine Beobachtung, die von Tauscher und Greenberg (Tauscher & Greenberg, 1997, S.406) sowie Cockburn und McKenzie bestätigt wird (Cockburn & McKenzie, 2001, S.915).

Adar, Teevan und Dumais analysieren Wiederbesuchsverhalten im Webkontext unter der Verwendung von drei Datenquellen: Suchmaschinen-Interaktionslogs, qualitativen Nutzerabsichten und Webseiteninhalten (Adar et al., 2008). Sie erkennen vier Wiederbesuchsmuster bzw. zwölf Verhaltenscluster, die jeweils eigene Verhaltens-, Inhalts- und Webseitenstrukturcharakteristika besitzen. Aufbauend auf diesen Charakteristika klassifizieren sie Webseiten nach raschen (<1 Std.), mittleren (>1 Std. und <24 Std.), langsamen (>24 Std.) und hybriden (< 1Std. und >24 Std.) Wiederbesuchsraten (Adar et al., 2008, S.7). Webseiten, die langsame Wiederbesuchsmuster aufweisen, sind nach Adar, Teevan und Dumais diejenigen, die auch tatsächlich wiedergefunden werden. Sie repräsentieren die Langzeitinteressen bzw. die unregelmäßigen Informationsbedürfnisse der Nutzer. Zeitlich gesehen ist eine Konzentration der Wiederbesuche nach einer Woche zu erkennen. Thematisch orientieren sich die Seiten Richtung Geschäftsangelegenheiten und Finanzen, Sportaktivitäten, Filme sowie Religion. Die Hauptstrategie zum Wiederfinden ist eine erneute Suchanfrage in Suchmaschinen (Adar et al., 2008, S.7). Wiederfindensverhalten mit Suchmaschinen wird im

Folgenden näher charakterisiert.

Wiederfinden mit Websuchmaschinen ist sehr speziell, da Websuchmaschinenergebnisse bzw. Webseiten im Allgemeinen — im Vergleich zum eigenen Dateisystem oder E-Mails — einen sehr dynamischen Informationsraum darstellen.

Eine der ersten Arbeiten, die Suchanfragemlogs zur Analyse des Wiederfindensverhalten analysiert, ist eine Studie von Teevan und Kollegen, die Re-finding-Suchanfragen von 114 Nutzern im Zeitraum eines Jahres mit der Yahoo-Websuchmaschine untersucht (Teevan et al., 2006). Sie unterstellen einem Nutzer dann Wiederfindensabsicht, wenn er zu unterschiedlichen Zeitpunkten auf das gleiche Suchmaschinenergebnis klickt, wobei die Suchanfragen per se dabei durchaus unterschiedlich sein können (Teevan et al., 2006, S.703). Sie stellen fest, dass bei 40% aller Suchanfragen Ergebnisse geklickt werden, die in früheren Sitzungen bereits angeklickt wurden. 71% der Wiederfindenssuchanfragen enthalten einen Term aus einer früheren Suchanfrage. In 13% führen identische Suchanfragen nicht zum Klick auf das gleiche Ergebnis (Teevan et al., 2006, S.703). Ferner stellen sie fest, dass Wiederfindensverhalten mit hoher Genauigkeit vorhergesagt werden kann, was vor allem für navigationsorientierte Suchanfragen gilt, bei denen eine Genauigkeit von 96% erreicht werden kann. Ein Wechsel im Ranking der Ergebnisse führt dazu, dass nur bei 53% der Suchanfragen ein Re-Klick auf ein Ergebnis erfolgt (Teevan et al., 2006, S.704). Im Folgejahr vertiefen Teevan und Kollegen ihre Analysen zum Wiederfindensverhalten mit Websuchmaschinen (Teevan, Adar et al., 2007). Dabei untersuchen sie unter anderem, welche lexikalischen Veränderungen Suchanfragen mit Wiederfindensabsicht durchleben, wobei häufig nur einzelne Worte weggelassen werden oder deren Reihenfolge getauscht wird (Teevan, Adar et al., 2007, S.154). Ferner weiten sie ihre Analyse zu Einflussfaktoren auf erneute Klicks aus. Neben der Tatsache, dass eine Veränderung im Ranking die Wahrscheinlichkeit eines erneuten Klicks verringert, fördern ihre Analysen auch zu Tage, dass die Veränderung im Ranking zusätzlich das Wiederfinden verlangsamt, wenn Ergebnisse weiter nach unten in der Liste rutschen, aber auch beschleunigen, wenn Ergebnisse weiter nach oben steigen (Teevan, Adar et al., 2007, S.156). In einer späteren Studie untersucht Teevan, welche Faktoren Einfluss darauf haben, ob ein Suchmaschinenergebnis in der SERP erinnert wird (Teevan, 2007). Entscheidende Faktoren für das Erinnern eines SERP-Ergebnisses sind: Die Position des Ergebnisses (je weiter oben das Ergebnis gerankt war, desto wahrscheinlicher wurde es erinnert) sowie ob das Ergebnis geklickt wurde (Teevan, 2007, S.25).

Sanderson und Dumais bestätigen ebenfalls das Auftreten von Re-finding-Verhalten in Querylogs (Sanderson & Dumais, 2007). Sie fokussieren dabei stärker auf die zeitlichen Charakteristika von wiederholten Suchanfragen und zeigen, dass navigationsorientierte Suchanfragen über längere Zeiträume auftreten als nicht-navigationsorientierte Suchanfragen (Sanderson & Dumais, 2007).

Tyler und Teevan analysieren das Wiederfindensverhalten sowohl innerhalb einer Session als auch über mehrere Sessions hinweg (*session-level vs. cross-session re-finding*). Als Datengrundlage verfügen sie über Querylogs, gecrawlten Webseiteninhalten, sowie Webbrowserlogs (Tyler & Teevan, 2010). In Bezug auf die Evolution von Suchanfragen zeigen sie, dass initiale Suchanfragen oft wesentlich schlechter sind, als ihr Re-finding-Pendant, wobei die Veränderungen, die an der Query vorgenommen werden umso stärker sind, je weiter die initiale Suche in der Vergangenheit zurück liegt (Tyler & Teevan, 2010, S.195). Inhaltlich beziehen sich Re-finding-Suchanfragen stärker auf den Webseiteninhalt, woraus man schließen kann, dass sich Nutzer an den Inhalt erinnern können (Tyler & Teevan, 2010, S.196). Die Querylänge stellt den größten Unterschied zwischen Intra- und Inter-Session-Suchanfragen dar, wobei Re-finding-Suchanfragen deutlich kürzer sind (Tyler & Teevan, 2010, S.195). In Bezug auf die Absichten der Nutzer interpretieren sie Wiederfinden innerhalb einer Sitzung als Re-Evaluierung, wohingegen sie Wiederfinden über mehrere Sitzungen hinweg als das Aufgreifen einer angefangenen Aufgabe charakterisieren (Tyler & Teevan, 2010, S.198). Letztendlich stellt sich aber generell die Frage, ob man bei Handlungen innerhalb einer

Session überhaupt von Re-finding-Handlungen sprechen kann.

Aktuelle Untersuchungen des Wiederfindensverhaltens mit Websuchmaschinen von Sadeghi et al. versuchen, mithilfe von maschinellen Lernverfahren Vorhersagen zu treffen, ob Nutzer die Absicht besitzen, Informationen wiederzufinden. Sie ermitteln auch, wie hoch der Schwierigkeitsgrad dieser Aufgabe ist (Sadeghi et al., 2015). Ein Novum in ihrer Herangehensweise stellt der Verzicht auf die Suchhistorie des Nutzers dar. Zur Vorhersage verwenden sie nur Features der aktuellen Sitzung (Sadeghi et al., 2015, S.724). Zusätzlich versuchen sie, dieses Problem auf Suchergebnisse von sog. *verticals*, also Bildersuche oder Nachrichtensuche, zu übertragen (Sadeghi et al., 2017).

Wie das Wiederfinden sowohl clientseitig, also durch zusätzliche Funktionalität oder UI-Features als auch algorithmisch unterstützt werden kann, wird in Abschnitt 8 vorgestellt. Diese Konzepte münden in eigenen Designvorschlägen für die Unterstützung des Wiederfindensverhaltens auf Twitter.

## 2.3. PIM und PIM-Praktiken: Ein Fazit

Die Darstellung der PIM-Forschungslandschaft in Bezug auf die Erforschung des Verhaltens mit einzelnen Informationstypen und in Bezug auf deren Hauptgebiete haben ein breites Spektrum an Erkenntnissen geliefert, die letztendlich die Frage aufwerfen, ob diese Erkenntnisse bzw. Charakteristika auch im Kontext des PIM-Verhaltens auf Twitter festgestellt werden können. Die Zusammenfassung der wichtigsten Punkte dient gleichzeitig als Basis für die Formulierung von Forschungsfragen in den Abschnitten 5.1 und 6.1.

In den Forschungsarbeiten werden zahlreiche Idiome oder Thesen zu Tage gefördert, die bei der Interpretation der Erkenntnisse aus den beiden Studien helfen und in Kapitel 7 zur Darstellung der wesentlichen Ergebnisse der Arbeit genutzt werden können.

- Wen prägt den Begriff des *post retrieval value*, also die nachträgliche Erkenntnis, dass ein verworfenes Informationsobjekt doch relevant gewesen wäre (Wen, 2003).
- Bruce formuliert die Personal-Anticipated-Information-Need-Hypothese, die besagt, dass Nutzer in Bezug auf Information stets mit einer Aufbewahrensentscheidung konfrontiert sind, die eine Vorhersage zukünftiger Informationsbedürfnisse — worin Personen aber Defizite zeigen — bzw. eine Kosten-Nutzen-Analyse erfordert (Bruce, 2005).
- Bergman und Whittaker stellen das Problem der *deferred evaluation* vor, das besagt, dass der Informationsüberfluss dazu beiträgt, dass Nutzer die Relevanzprüfung von Informationsobjekten auf einen späteren Zeitpunkt verschieben, da in der aktuellen Situation oft zu wenig Zeit zur Verfügung steht, um Informationen einer vollständigen Relevanzbewertung zu unterziehen. Als Konsequenz daraus werden noch nicht bewertete Informationen an Orten angehäuft. Diese werden jedoch selten bis nie besucht, da die Erinnerung an die noch zu tätigende Relevanzbewertung fehlt. Werden diese Orte nach längerer Zeit wieder besucht, kommt es anstatt eines Wiederfindens zu einem erneuten Entdecken (*Re-discovery*) (Bergman & Whittaker, 2016).
- Jones et al. stellen ein ähnliches Problem auch für bereits als relevant bewertete und aufbewahrte Information fest. Sie sprechen vom Out-of-sight-out-of-mind-Problem. Es beschreibt den Umstand, dass relevante Information archiviert wird und somit die Aufmerksamkeitsebene des Nutzers verlässt (Jones, Phuwanartnurak et al., 2005)
- Besondere Relevanz für das Wiederfinden von Information hat die von Jones beschriebene First-Impression-Hypothese, welche besagt, dass sich Nutzer an diejenigen Charakteristika

eines Informationsobjekts besonders gut erinnern können, die dazu benutzt wurden, um aus einer Menge an Informationen dieses herauszufiltern bzw. anhand dieser Charakteristika eine Priorisierung von Information vorzunehmen. Die erinnerten Charakteristika werden dann zum Wiederfinden genutzt (Jones, 2013).

- *Orienteering* vs. *teleporting*: In den Forschungsarbeiten zum Wiederfinden werden immer wieder die beiden Verhaltensweisen navigationsbasiertes und auf Schlüsselwortsuche basiertes Wiederfinden gegenübergestellt. Jedoch sind auch Mischformen zu erkennen.
- In verschiedenen Studien zum Wiederfinden konnten spezifische Charakteristika des Wiederfindensverhaltens festgestellt werden:
  - Alter der Informationsobjekte: Erst kürzlich gefundene Dokumente werden am häufigsten wiedergefunden (Fokus auf *recency*).
  - Häufigkeit des Wiederfindens: Es gibt einen Fokus auf Informationsobjekte, die immer wieder wiedergefunden werden.
  - Querystrategien: Bei der Verwendung des Teleporting-Ansatzes ist ein Fokus auf Suchanfragen zu erkennen, die Personennamen enthalten.

Die in den vorherigen Abschnitten erwähnten Einflussfaktoren auf Aufbewahren und Wiederfinden, wie z. B. Faktoren, die das Wiederfinden zu einer schwierigen, frustrierenden Aufgabe machen können, werden hier nicht nochmals explizit aufgeführt.





# 3

## **Social-Media-Forschung: Twitter und Social-Media-PIM im Fokus**

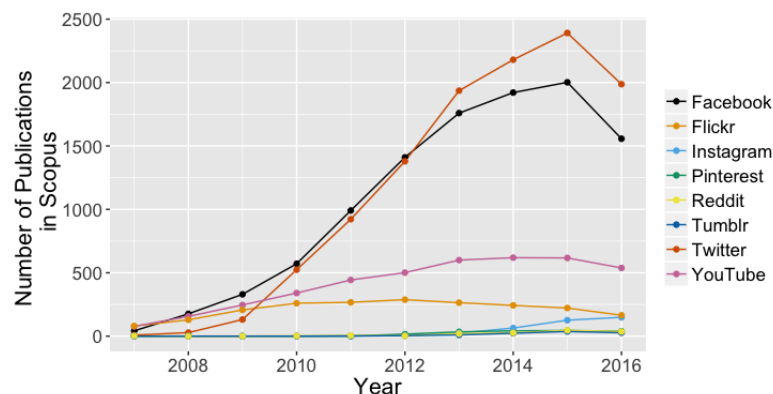
Kapitel 3 ist in drei Teile gegliedert. In Abschnitt 3.1 wird Twitter als Gegenstand der Forschung eingeführt. Es wird dargestellt, wie stark Twitter im Fokus der Wissenschaft steht, wobei die Heterogenität der Twitter-Forschung anhand von exemplarisch ausgewählten Forschungsarbeiten aus unterschiedlichen Disziplinen verdeutlicht wird. Abschnitt 3.2 stellt relevante Literatur im Kontext von Social Media und Informationsverhalten vor. Entlang der PIM-Dimensionen — Finden, Aufbewahren und Wiederfinden — werden für das Forschungsvorhaben relevante Arbeiten vorgestellt. Der Fokus liegt hauptsächlich auf Twitter-Forschung, jedoch wird der Literaturüberblick auch auf andere Social-Media-Plattformen — wie Facebook, Instagram oder Pinterest — ausgeweitet, so dass höchst einschlägige Arbeiten, die das PIM-Nutzerverhalten auf anderen Plattformen analysieren, an den entsprechenden Stellen besprochen werden. Das Kapitel schließt in Abschnitt 3.3 mit einer Zusammenfassung der wichtigsten Aspekte der relevanten Literatur, welche schließlich als Grundlage für die Formulierung der Leitfragen der beiden Studien dienen. Im Sinne der Cooper-Taxonomie, welche Literatur-Reviews entlang der Dimensionen Fokus, Ziel, Perspektive und Grad der Abdeckung einteilt, handelt es sich in diesem Kapitel um ein Review, das auf Ergebnisse fokussiert. Ziel des Reviews ist es, eine Generalisierung durchzuführen, also existierende Literatur zusammenzufassen und zentrale Aussagen über das Forschungsfeld zu treffen (Cooper, 1988, S.109). Dabei wird eine neutrale Perspektive eingenommen. Der Grad der Abdeckung ist so gewählt, dass es die zentrale Literatur übersieht, aber keineswegs auf Vollständigkeit abzielt (Cooper, 1988, S.109).

Einige Arbeiten die in den kommenden Abschnitten vorgestellt werden, wurden erst nach dem Forschungsbeginn veröffentlicht und haben daher nicht explizit auf die Konzeption der Studien Einfluss genommen. Die Ergebnisse dieser Arbeiten werden jedoch auch reflektiert und diskutiert.

### **3.1. Twitter-Forschung im Überblick**

Keine Social-Media-Applikation steht so stark im Fokus der Forschung wie Twitter. Diese Aussage lässt sich auch anhand von Daten verifizieren. Abbildung 3.1 zeigt einen Überblick über die An-

zahl an Veröffentlichungen (Beiträge in Konferenzbänden und Zeitschriftenartikel) pro Jahr in den letzten 10 Jahren zu bekannten Social-Media-Applikationen in der Literaturdatenbank Scopus. Für Twitter und Facebook ist ein stetiger Zuwachs zu erkennen, was bedeutet, dass jedes Jahr mehr Artikel veröffentlicht wurden als im Vorjahr. Ab dem Jahr 2013 beginnt Twitter Facebook als die am häufigsten untersuchte Plattform abzulösen. Vermutlich werden auch im Jahr 2016 wieder mehr Publikationen zu erwarten sein als im Vorjahr, jedoch waren diese zum Untersuchungszeitpunkt noch nicht in Scopus vermerkt, wodurch sich eine fallende Entwicklung ergibt.



**Abbildung 3.1.:** Anzahl an Veröffentlichungen (pro Jahr) zu bekannten Social-Media-Applikationen in der Literaturdatenbank Scopus

Betrachtet man andere Social-Media-Anwendungen, die wesentlich älter sind als Twitter, wie YouTube oder Reddit, so ist es doch erstaunlich, wie deutlich der Unterschied in Hinblick auf die Veröffentlichungszahlen ist. Es gibt praktisch keine Disziplin, die nicht auf eine Art und Weise Twitter-Daten für ihre Forschungszwecke nutzbar macht. Gründe hierfür können die generelle Öffentlichkeit Twitters sein, da man auch Tweets lesen kann, ohne einen Account bei der Plattform haben zu müssen, und der einfache Zugang zu Twitter-Daten via Twitter-API. Die Veröffentlichungsstatistik zu Twitter zeigt, dass ein holistisches Literatur-Review sämtlicher Twitter-Arbeiten basierend auf vorgegebenen Standards — siehe z. B. die Arbeiten von Schryen, Wagner und Belian (Schryen et al., 2015) — in dieser Arbeit nicht geleistet werden kann. Im Folgenden wird durch frühe Kategorisierungsansätze und Literatur-Surveys, sowie exemplarischer Nennung von Studien aus den Sozialwissenschaften, das breite Spektrum der Twitter-Forschung vorgestellt.

Frühe Studien zu Twitter nehmen grobe Klassifizierungen von Twitter-Forschung vor. Barnes und Böhringer klassifizieren Twitter-Forschung in zwei Kategorien: *Micro-blogging Verstehen* (beschreibende Statistiken zu Twitter) und *Micro-blogging* in speziellen Anwendungssituationen (z. B. im Unternehmen) (Barnes & Böhringer, 2011). Dann nennt explizit Disziplinen wie Gesundheitswesen, Politik, Regierung, Unternehmen, Erziehung und Lernen, Journalismus, Augenzeugenberichte von Nachrichten (Dann, 2010). Zahlreiche Meta-Studien versuchen die Dimensionen der Twitter-Forschung noch detaillierter und tiefer zu erfassen (R. Rogers, 2013; Williams et al., 2013b,a; Zimmer & Proferes, 2014). Rogers nimmt eine Meta-Analyse von Twitter-Studien aus medienwissenschaftlicher Perspektive vor (R. Rogers, 2013). Er stellt fest, dass Twitter-Forschung in den ersten Jahren vor allem durch die öffentliche Wahrnehmung der Plattform geprägt ist und teilt die Forschung entlang folgender drei Phasen ein:

1. Twitter Studies I: *Banal, Phatic, Shallow*
2. Twitter Studies II: *Towards a news medium for event-following*
3. Twitter Studies III: *Twitter as (archived) data set*

Der ersten Phase rechnet er vor allem Studien zu, die sich mit Tweet-Inhalten auseinandersetzen, welche in der öffentlichen Wahrnehmung lediglich als banales, „pointless babble“ oder „daily chatter“ gesehen werden, bei denen der Fokus auf dem Aktuellen und Schnelllebigen liegt, das keinen nachhaltigen Wert besitzt (R. Rogers, 2013). Er führt hier bekannte Studien an wie von Java et al., Naaman, Boase und Lai oder Marwick und boyd, die alle eine qualitative Inhaltsanalyse von Tweets vornehmen und Tweet-Inhaltskategorien sowie Nutzerklassen ausarbeiten (Java et al., 2007; Naaman et al., 2010; Marwick & boyd, 2011). Naaman und Kollegen entwickeln mit der Unterscheidung zwischen *meformers* — Nutzer die hauptsächlich Nachrichten über sich selbst versenden — und *informers* — Nutzer die hauptsächlich Informationen senden bzw. weiterleiten — die wohl bekanntesten Twitter-Nutzerklassen (Naaman et al., 2010). Die zweite Phase der Twitter-Forschung ist geprägt von der Nutzung Twitters während öffentlicher Ereignisse wie den Protesten im Iran, oder dem Erdbeben in Haiti sowie dem mobilisierenden Charakter Twitters, der durch die Konversationsmöglichkeiten der in diesen Jahren eingeführten Features *@Mention* und dem Hashtag entsteht (Honey & Herring, 2009; Muralidharan et al., 2011; Gaffney, 2010). Die dritte Phase ist geprägt von der Anerkennung Twitters als Forschungsobjekt und der Wahrnehmung von Tweets als kulturelle Artefakte, die sogar so viel Wert besitzen, um Kollektionen von ihnen zu erstellen. Zu dieser Zeit werden Ansätze von Bibliotheken entwickelt (z. B. von der *Library of Congress*), um Tweets nachhaltig aufbewahren zu lassen und ernsthafte Forschung zu betreiben, wie die Vorhersage der Finanzmarktentwicklung oder Informationen über das Gesundheitswesen (Bollen et al., 2010; Collier et al., 2011).

Williams, Terras und Warwick nehmen das Resultat einer Literatursuche bei den beiden Datenbanken Scopus und Web of Science als Ausgangspunkt für eine Meta-Analyse (Williams et al., 2013b). Für den Suchzeitraum 2007 bis 2011 finden sie 1161 Arbeiten, die den Term Twitter enthalten, wobei aber nur 575 tatsächlich relevant sind und sich auf die Social-Media-Plattform beziehen. Diese Artikel klassifizieren sie entlang von drei Dimensionen: dem Analysefokus (Tweetinhalt, Nutzer, Technologie und Konzeptpapiere), der Methode, die genutzt wird (Analytik z. B. Inhaltsanalyse, Design und Entwicklung, Literaturübersichten, *knowledge discovery* z. B. mit Data-Mining- oder Text-Mining-Verfahren) und der Domänenperspektive (sie identifizieren 13 Domänen u.a. Gesundheit, Bibliothekswissenschaft, IT-Sicherheit etc.). In ihren Analysen stellen sie fest, dass sich 80% der untersuchten Veröffentlichungen auf den Tweetinhalt oder den Nutzer konzentrieren. Der Großteil der Studien verwendet Data-Mining-Techniken, um ein Verständnis von einer großen Menge an Tweets zu erlangen, was zeigt, dass Twitter-Studien hauptsächlich als quantitative Forschung zu interpretieren sind.

Williams, Terras und Warwick wenden dasselbe Verfahren abermals im Kontext der Fachdatenbank *PubMed*<sup>1</sup> an, um zu identifizieren, wie die Medizin, Twitter in der Forschung nutzt (Williams et al., 2013a). Von 123 Veröffentlichungen, die Twitter als Term im Titel oder Abstract enthalten, sind nur 30 tatsächlich relevant. Sie verwenden dieselben Dimensionen wie in ihrer früheren Arbeit. Als Subdomänen identifizieren sie Studien zu Psychologie, Influenza-Forschung oder Neurologie. Ein Beispiel ist die Untersuchung von Tweets, die Nutzer zu ihrem Gesundheitszustand während des H1N1-Ausbruchs 2009 versenden (Collier et al., 2011). Aber auch die Nutzbarmachung von Twitter in der Ärzte- oder Apothekerausbildung wird in Forschungsarbeiten thematisiert (Fox & Varadarajan, 2011).

Die bis dato aktuellste Meta-Studie führen Zimmer und Proferes mit einer *Topology of Twitter re-*

<sup>1</sup><https://www.ncbi.nlm.nih.gov/pubmed>

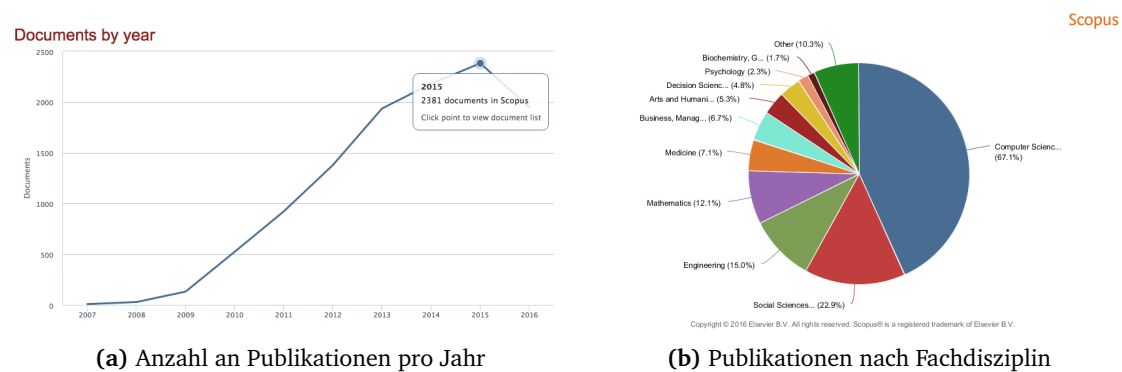
Datenbank	Einträge	Suchanfrage	Ergebnis
The ACM Full-Text Collection	~ 450 000	Twitter	2554
The ACM Guide to Computing Literature	~ 2,5 Mio.	Twitter	5280
IEEE Xplore Digital Library	~ 4 Mio.	Twitter	3425
Web of Science All Databases	~ 1 Mrd.	Twitter	3339
Wiley Online Library Journals only	~ 6 Mio.	Twitter	9382
Scopus Conference Paper and Article	~ 57 Mio.	TITLE-ABS-KEY ( twitter ) AND PUBYEAR > 2006 AND (LIMIT-TO ( DOCTYPE , cp ) OR LIMIT-TO ( DOCTYPE , ar ) )	11462

**Tabelle 3.1.:** Suche in bibliographischen Datenbanken nach dem Term Twitter (Stand: November 2016)

search durch (Zimmer & Proferes, 2014). Bei der Inhaltsanalyse der Volltexte von 382 Veröffentlichungen aus dem Zeitraum 2006 bis 2012 fokussieren sie auf die Disziplinen, in denen Twitter-Daten genutzt werden, die Methoden, mit denen Twitter-Daten erhoben und analysiert werden, sowie die ethischen Fragestellungen, die bei der Forschung mit Twitter-Daten auftreten (Zimmer & Proferes, 2014). Auch sie stellen eine starke disziplinäre Heterogenität fest. Neben Fächern wie Informatik, Informationswissenschaft und Kommunikationswissenschaft stellen sie ein verstärktes Interesse an Twitter-Studien seitens der Fächer Politikwissenschaft und Soziologie fest. Inhaltlich gesehen identifizieren sie die Analyse von Tweet-Verbreitungsmustern, Follower-Followee-Netzwerken, Sentiment-Analysen sowie die Nutzung Twitters von bestimmten Nutzergruppen als Hauptinteressenlagen dieser Studien. Die Untersuchung von Zimmer und Proferes besitzt als Alleinstellungsmerkmal die Analyse der Größe der Tweet-Korpora, die in den Studien verwendet werden, sowie deren Quellen. Ihre Beobachtung, dass mehr als die Hälfte (etwa 60%) aller Arbeiten Tweets analysieren, die via API gesammelt wurden, ist nicht weiter verwunderlich, stützt jedoch die getroffene Annahme, dass ein Grund für die exzessive Twitter-Forschung die Einfachheit der Datenakquise ist (Zimmer & Proferes, 2014, S.256). Sie resümieren, dass sowohl in Bezug auf Disziplinen als auch den verwendeten Methoden zur Datenerhebung und -analyse, ein kontinuierlich zunehmender Trend zu erkennen ist, verweisen aber auch auf die sich herausbildenden ethischen Herausforderungen, die mit dieser Art Forschung einhergehen und plädieren für die Erstellung von ethischen Richtlinien zur Twitter-Forschung (Zimmer & Proferes, 2014, S.258).

Die letzten Abschnitte haben gezeigt, dass es durchaus schon Bestrebungen gibt, die versuchen, die Dimensionen der Twitter-Forschung zu erfassen und Klassifikationsansätze für die von sehr starker Heterogenität geprägten Forschungslandschaft erstellen. Letztendlich müssen sich aber alle Ansätze der Kritik stellen, dass sie nur einen Bruchteil der existierenden Publikationen in ihre Meta-Analysen miteinbeziehen. Interessanterweise gibt es noch keine Ansätze, die automatische Verfahren zur Themen-Erkennung (z. B. *topic modelling* via *latent dirichlet allocation* (LDA)) unter Verwendung eines größeren Korpus aus Titel und Abstracts anwenden, wie dies in anderen Kontexten z.B. zur Erkundung der Themengebiete in der informationswissenschaftlichen Forschungslandschaft bereits getan wurde (Sugimoto et al., 2011). Jedoch würde auch hier nur eine thematische Gliederung vorgenommen werden. Ähnlich wie es die Arbeiten in den Abschnitten weiter oben zeigen, könnte zusätzlich eine Klassifizierung hinsichtlich der Methodik (qualitativ, quantitativ, experimentell, inhaltsanalytisch) oder des Untersuchungsgegenstands (Tweetinhalt, Nutzer, Plattform) vorgenommen werden. Hier wäre sicherlich Raum für weitere Forschungsarbeiten.

Tabelle 3.1 zeigt, dass, obwohl Twitter im März 2016 gerade mal 10 Jahre alt geworden ist, bereits mehrere Tausend Publikationen veröffentlicht wurden, die sich dem Microblogging-Dienst zuwenden (Twitter, 2016b). Zum einen zeigen die Zahlen, dass eine ganzheitliche Darstellung der Twitter-Forschung nicht möglich ist. Zum anderen sind sie jedoch mit Vorsicht zu genießen, da



**Abbildung 3.2.:** Publikationen zum Thema Twitter aufgeschlüsselt nach Veröffentlichungsjahr und Fachgebiet (Export von: <https://www.scopus.com/>)

| Stand: November 2016)

nicht alle Arbeiten, die den Term Twitter im Abstract tragen, auch tatsächlich Arbeiten sind, die Twitter oder Twitter-Daten als Untersuchungsobjekt besitzen, wie auch schon Zimmer und Proferes anmerken (Zimmer & Proferes, 2014). Diese Tatsache lässt sich auch anhand der Ergebnisse des *The ACM Guide to Computing Literature* erkennen, da der Anstieg im Vergleich zu den Arbeiten in *The ACM Full-Text Collection* vor allem auf populärwissenschaftliche Arbeiten und Ratgeber (z. B. *Twitter for dummies* (Fitton et al., 2010)) zurückzuführen ist. Besonders das Ergebnis der *Wiley Online Library* enthält vermutlich sehr viele falsch-positive Ergebnisse, da dort auch Artikel aus der *Zeitschrift für Tierpsychologie* unter den Ergebnissen zu finden sind, was vermutlich auf das Zwischern (*to twitter*) realer Vögel zurückzuführen ist. Eine Suche begrenzt auf Journal-Auswahl ist dort leider nicht möglich.

Führt man eine detaillierte Analyse der Einträge in der Literaturdatenbank Scopus durch, wird eine Entwicklung der Publikationen über die letzten Jahre sowie eine Aufteilung nach Fachgebieten nachvollziehbar. Hierbei wurde die Suche auf den Zeitraum ab 2007 bis heute begrenzt, sowie nur *Conference Proceedings* und *Journal Article* als Quellen ausgewählt. Dies führt zu einer Gesamtzahl von 11462 Veröffentlichungen. Abbildung 3.2a zeigt den kontinuierlichen Anstieg von Publikationen zu Twitter über die letzten Jahre bis zu einem Höhepunkt von 2381 Veröffentlichungen alleine im Jahr 2015. Abbildung 3.2b lässt erkennen, dass mit 67,1% die meisten Publikationen aus dem Informatikkontext stammen, gefolgt von 22,9% aus dem Fachgebiet *Social Science*. Eine stichprobenartige Untersuchung der Publikationen aus den Fachbereichen Mathematik und Ingenieurwissenschaften lässt vermuten, dass diese auch dem Informatikkontext zugeordnet werden könnten. Die Aufteilung der Fachgebiete ist sicherlich diskutierbar, da Fächer wie die Psychologie oder die Wirtschaftswissenschaften (hier *Business, Management*) traditionell eher als Sub-Disziplinen der Sozialwissenschaften gesehen werden, bei Scopus aber als eigene Themengebiete aufgeführt sind. Im Folgenden wird anhand des Themengebiets *Social Sciences* die Vielfalt der Twitter-Forschung deutlich gemacht. Forschungsarbeiten im Informatikkontext sind ohnehin einschlägig für diese Arbeit und werden gesondert in Abschnitt 3.2 beschrieben.

#### Twitter-Forschung mit Social-Sciences-Bezug

Im Kontext sozialwissenschaftlicher Arbeiten wird deutlich, dass Tweets oder Account-Beschreibungen als Proxy für Aussagen genutzt werden, die sonst über Fragebögen oder Interviews gesammelt werden bzw. werden müssten. Diese Form der Datengewinnung ist vor allem in der Subdisziplin der *Computational Social Science* prävalent, deren Veröffentlichungen z. B. auf Konferenzen wie der *International Conference on Web and Social Media* (ICWSM) zu finden sind.

Ein breites Spektrum an Studien lässt sich im Kontext von sozio-politischen Ereignissen (oft wird von Revolutionen gesprochen (Büttner & Büttner, 2016, S.2208)) und der Charakterisierung von Twitter-Nutzung und Tweet-Inhalten während Protesten erkennen (Büttner & Büttner, 2016). Beispiele sind Untersuchungen zu den Protesten in Ägypten 2011 oder zur Occupy-Wallstreet-Bewegung (Bruns et al., 2013; Gleason, 2013). Weitere Arbeiten im politikwissenschaftlichen Kontext sind die Vorhersage von Wahlergebnissen oder die Nutzung von Twitter-Daten und Tweet-Sendeverhalten zur Analyse von Wahlkämpfen und Verhalten am Wahltag (Tumasjan et al., 2010). Dabei kann keine Beschränkung auf bestimmte geographische Regionen ausgemacht werden. So finden sich derartige Studien für unterschiedliche Länder wie Singapur (Skoric et al., 2012), Kanada (Gruzd & Roy, 2014) oder Schweden (Larsson & Moe, 2012). Weitere Studien erforschen die Konversationscharakteristika von Politikern und politischen Parteien (Lietz et al., 2014; Borondo et al., 2014) oder wie politische Institutionen, wie der *United States Congress*, Twitter nutzen (Golbeck et al., 2010). Zudem werden Experimente durchgeführt, die die politische Gesinnung von Personen über ihre Tweet-Inhalte vorherzusagen versuchen (Colleoni et al., 2014). Im Lichte der Abstimmung zum Austritt Großbritanniens aus der EU (*Brexit*) und den Präsidentschaftswahlen in den USA, die beide eher ein unerwartetes Ende genommen haben und in beiden Fällen der Meinungsbildungsprozess sehr stark durch Social Media beeinflusst war, sind Studien von Interesse, die die Meinungsbildung via Bot-Accounts auf Twitter untersuchen (Howard & Kollanyi, 2016; Kollanyi et al., 2016). Betrachtet man speziell soziologisch/psychologische Studien, die auf den Nutzer fokussieren, sind Studien zu erwähnen, die auf der Basis von Tweets und Account-Beschreibungen die Unterschiede zwischen Twitter-Nutzung in Stadt- und Land-Regionen (Hecht & Stephens, 2014) oder zwischen den Geschlechtern (Cunha et al., 2012) analysieren. Generell werden auch menschliche Alltagssituationen untersucht, wie der Beginn und das Ende von zwischenmenschlichen Beziehungen (Garimella et al., 2014) oder Ernährungsmuster in der Bevölkerung (Abbar et al., 2015). Hossain et al. versuchen mit maschinellen Lernverfahren Muster in Tweets zu erkennen, die unter Alkoholeinfluss gesendet wurden, um Nutzer zu identifizieren, die häufig betrunken twittern (Hossain et al., 2016). Auch Gefühle, Emotionen oder psychische Störungen von Nutzern wie Einsamkeit (Kivran-Swaine et al., 2014) oder Depression (Choudhury et al., 2013) werden auf der Basis von Twitter-Daten analysiert. Ein breites Themengebiet stellt auch die Nutzung von twitterähnlichen Microblog-Plattformen in Unternehmen dar (J. Zhang et al., 2010).

Im Kontext medienwissenschaftlicher Studien können Arbeiten angeführt werden, die die Themen Twitter und *Social TV* miteinander verknüpfen und Twitter als Erweiterung von klassischen Medien wie dem Fernsehen analysieren (Burghardt et al., 2013). Ferner werden Konversationen auf Twitter, die große TV-Ereignisse begleiten, untersucht (Shamma et al., 2010). All diese Untersuchungen sind unter dem Oberbegriff der *Second Screen Interaction* zusammenfassbar (Mukherjee & Jansen, 2016).

Social-Media-Plattformen werden in den meisten Studien zwar häufig isoliert betrachtet, dennoch gibt es auch Untersuchungen, die Unterschiede in den Inhalten und Nutzungsweisen von verschiedenen Plattformen analysieren und somit Cross-Plattform-Studien durchführen (Ottoni et al., 2014; Lim et al., 2015; Mukherjee & Jansen, 2016). Diese Studien versuchen der Tatsache Rechnung zu tragen, dass Nutzer häufig nicht nur eine Plattform verwenden, sondern mehrere Accounts auf unterschiedlichen Plattformen besitzen und dieselben Inhalte auch auf diesen unterschiedlichen Plattformen verbreiten, oder verschiedene Plattformen für unterschiedliche Zwecke nutzen. Ein aktuelles Beispiel stellt die Arbeit von Zhao, Lampe und Ellison dar, welche die Social-Media-Ökologie der Nutzer analysieren, also die wechselseitige Nutzung verschiedenster Social-Media-Plattformen und dabei die mentalen Modelle, die hinter der Nutzung jeder einzelner Plattform stehen, zu ergründen versuchen (Zhao et al., 2016).

## 3.2. Social-Media-Informationsverhalten mit Fokus auf PIM-Praktiken

Im Folgenden wird eine detaillierte Analyse und Darstellung von Forschungsarbeiten im Kontext von Informationsverhalten auf Social Media entlang der PIM-Dimensionen vorgenommen. Diese Verhaltensweisen sind: Finden (*finding*), Management (als Zusammenfassung aus *keeping*, *organizing* und *maintaining*) und Re-finding. Auch hier liegt der Fokus hauptsächlich auf Twitter, jedoch werden zum Teil auch Arbeiten zu anderen Plattformen herangezogen, falls sie für die vorliegende PIM-Praktik bzw. Forschungsfrage relevant sind. Dabei wird es interessant sein zu sehen, dass trotz der Tatsache, dass Twitter bereits exzessiv beforscht wurde, noch keine oder wenige Erkenntnisse über das PIM-Verhalten der Nutzer auf dieser Plattform existieren.

### 3.2.1. Social-Media-Inhalte Finden

Forschungsarbeiten, die das Auffinden (*finding*) von relevanten Social-Media-Nachrichten zum Inhalt haben, können thematisch in folgende Teilbereiche untergliedert werden:

- Gerichtete Suche, Relevanzkriterien und (Re-)Ranking von Tweets
- Social Media als soziale Frage-Antwort-Systeme
- Informationsdiffusion, RT-Vorhersage und *serendipity*
- Informationserschließung und Hashtagnutzung

Diese Dimensionen werden nun im einzelnen durch relevante Studien charakterisiert.

#### 3.2.1.1. Gerichtete Suche, Relevanzkriterien und Microblogsuche

Zentral für die vorliegenden Studien sind Arbeiten, welche Informationsbedürfnisse und Informationssuchverhalten auf Social-Media-Plattformen untersuchen. Teevan, Ramage und Ringel Morris untersuchen die Nutzungsscharakteristika der Twitter-Suche, indem sie Suchanfrage-Logs auswerten (Teevan et al., 2011). Als Vorstudie, um Gründe für die Nutzung der Twitter-Suche zu eruieren, führen sie eine qualitative Umfrage durch. Drei Motive für die Twitter-Suche werden dabei evident (Teevan et al., 2011, S.37): Erstens, die Suche nach zeitkritischen bzw. Echtzeit-Informationen wie Nachrichten, die während Ereignissen entstehen und diese begleiten. Zweitens, die Suche nach sozialen Informationen, z. B. die Suche nach neuen Accounts oder Tweets von bestimmten Nutzern. Drittens, die Suche nach bestimmten Themen, wobei diese selbst eher zeitkritische oder soziale Schwerpunkte besitzen. In der quantitativen Analyse der Twitter-Suchanfragen vergleichen sie diese mit einer Stichprobe an Suchanfragelogs einer klassischen Suchmaschine. Folgende Hauptunterschiede können ausgemacht werden (Teevan et al., 2011, S.38ff.):

- Twitter-Suchanfragen sind wesentlich kürzer, bestehen aus weniger Wörtern und enthalten Twitter-spezifische Zeichen für *@Mentions* oder *#* für Hashtags. Zudem werden auf Twitter häufig die Namen von Berühmtheiten gesucht, wohingegen in der Websuche hauptsächlich navigationsorientierte Suchanfragen auftreten.
- Twitter-Suchsitzungen sind wesentlich kürzer als Suchsitzungen in der Websuche und bestehen aus weniger Suchanfragen, wobei auch weniger einzigartige Suchanfragen vorkommen.
- In Bezug auf die Ergebnisse beider IR-Systeme stellen sie fest, dass die Sprache von Tweets, im Vergleich zur Sprache von Ergebnissen einer Websuchmaschinen-SERP, sehr unterschied-

lich ist. Darüber hinaus sind die Ergebnisse einer Websuchmaschine thematisch wesentlich breiter gefächert.

Ein Nachteil ihrer Studie ist die Tatsache, dass via der von ihnen gesammelten Logs keine Unterscheidung getroffen werden kann, ob die Suchanfragen tatsächlich vom Nutzer eingegeben wurden, oder der Nutzer nur auf einen Link (z. B. *trending topics*) geklickt hat (Teevan et al., 2011, S.39). In ihren Designvorschlägen argumentieren sie für eine bessere Unterstützung entlang der Dimensionen zeitkritische Suche, Suche nach Personen (z. B. durch Einbettung von Websuchvorschlägen), Hashtag-Erweiterung der Suche, Suchgeschichte des Nutzers und Suchanfragedisambiguierungsstrategien.

Vergleichbar mit dem qualitativen Teil der Studie von Teevan et al. ist die Studie von Elweiler und Harvey, die untersucht, wie Art und Charakteristika von Aufgaben und Informationsbedarf die Informationssuche mit/auf Twitter motivieren (Elweiler & Harvey, 2015). Hierfür verwenden sie zwei qualitative Datenquellen, nämlich Informationsbedürfnisse, die von Probanden via einer Tagebuchstudie kommuniziert wurden und Tweets, die Beschreibungen des Suchverhaltens auf Twitter enthalten. Elweiler und Harvey finden 8 Hauptgründe und 17 Variationen von Informationsbedürfnissen, die die Nutzung der Twitter-Suche motivieren. Auch bei ihnen zählen die Suche nach zeitkritischen, aktuellen Informationen, die Suche nach Personen und das Beobachten von Themen zu den Hauptbeweggründen (Elweiler & Harvey, 2015, S.269). Ein weiterer Hauptgrund, den sie identifizieren, ist das Wiederfinden von Tweets, welcher in Abschnitt 3.2.3 im Detail beschrieben wird. Auch wenn Informationssuchverhalten in Studien zu anderen Plattformen nicht explizit Untersuchungsgegenstand ist, so lassen sich dennoch vereinzelt Erkenntnisse hierzu finden. Auf Pinterest identifizieren Linder, Snodgrass und Kerne zwei Arten von Informationssuchverhalten: das *casual browsing* ohne spezifischen Informationsbedarf oder einem Ziel im Blick und die gerichtete Suche nach konkreten Informationen (Linder et al., 2014).

Neben der Identifikation von Informationsbedürfnissen und der Beschreibung des Suchverhaltens bei der Twitter-Suche hat es in der IR-Community in den letzten fünf bis acht Jahren zahlreiche Forschungsanstrengungen gegeben, um Microblogsuche zu verbessern (Efron, 2011). Mit ein Grund für diese Entwicklung ist sicherlich die Tatsache, dass Standardsuchmaschinen wie Google, Bing oder Yahoo damit begonnen haben, Tweets als Ergebnisse einer sog. Vertical-Suchmaschine mit in die SERP einzubauen. Vorangetrieben werden diese Anstrengungen auch durch den von der *National Institute of Standards and Technology* (NIST) im Rahmen der *Text REtrieval Conference* (TREC) etablierten *Microblog Track*, welcher seit 2011 jährlich stattfindet und eine Evaluierung von Retrieval-Strategien bzw. Modellen zur Microblogsuche in einem standardisierten Umfeld erlaubt (Ounis et al., 2011). Diese wurde anfangs noch mit tatsächlichen Testkollektionen durchgeführt (McCreadie et al., 2012), ist mittlerweile aber an den Charakter von Microblogs als *information stream* angepasst, so dass eine neue Variante via *TREC Microblog API* angeboten wird, die als *Evaluation as a service* (EaaS) charakterisiert werden kann (J. Lin & Efron, 2013; Rao et al., 2015). Der Aufbau von Testkollektionen für die Evaluation von Microblogsuche bleibt aber ein bestehendes Problem (Berendsen et al., 2013).

Aufgrund der besonderen Beschaffenheit von Tweets (Beschränkung auf 140 Zeichen, besondere Zeichen mit *@Mentions* und Hashtags, Verlinkungen innerhalb des Textes etc.) untersuchen zahlreiche Studien die Faktoren, die zur Interessanztheit eines Tweets beitragen, um daraus Features abzuleiten, die in einem Retrieval-Modell genutzt werden können. Motiviert sind diese Studien durch die Tatsache, dass klassische Verfahren, die mit Termstatistiken arbeiten (z. B. TFxIDF), aufgrund eben genannter Charakteristika und der daraus resultierenden Seltenheit von Termen (*sparsity*), keine optimalen Ergebnisse liefern können (Naveed et al., 2011). Vergleichbar mit dem quantitativen Teil der Studie von Teevan et al. ist die Untersuchung von Lin und Mishne, welche den *churn*, also die auffällig schnelle Veränderung von relevanten Themen und Suchanfragen bei



der Echtzeit-Suche analysieren (J. Lin & Mishne, 2012). Ziel ihrer Untersuchungen ist die Klärung der Frage, wie Echtzeit-Suche mit dem Problem der Termseltenheit umzugehen hat, wenn die zugrundeliegenden Verteilungen der Terme sich rapide ändern. Dies ist dann der Fall, wenn Themen für eine kurze Zeitspanne extrem relevant sind, aber schließlich das Rampenlicht verlassen (J. Lin & Mishne, 2012, S.503).

Im Fokus von Features für Retrieval-Modelle steht immer wieder die individuelle Interessenlage eines Tweets für den Nutzer und die Suche nach einer Art kleinster gemeinsamer Nenner, welche als Feature aus solchen Studien herausgearbeitet werden können, um eine bessere Personalisierung bei der Suche vornehmen zu können (André et al., 2012; Alonso et al., 2013). Zur Klärung der Frage, welche Charakteristika zur Interessantheit eines Tweets beitragen, wird eine Bandbreite an unterschiedlichen Methoden verwendet, die von großangelegten API-Datenanalysen über Crowdsourcing-Experimenten bis hin zu Eye-Tracking-Experimenten reichen (Alonso et al., 2013; Counts & Fisher, 2011). Counts und Fisher untersuchen in einer Eye-Tracking-Studie, wie Nutzer ihre Twitter-Timeline wahrnehmen (Counts & Fisher, 2011). Sie motivieren ihre Studie mit dem Argument, dass Twitter dem Nutzer das Verarbeiten von Information äußerst erschwert und ihm eine schnelle Relevanzentscheidung abverlangt. Dies kann aufgrund der hohen Dynamik der Plattform starke Konsequenzen haben, da vermeintlich relevante Informationen verloren gehen können (Counts & Fisher, 2011). Für sie ist der Interessantheitsgrad eines Tweets das entscheidende Kriterium, ob ein Tweet verarbeitet wird. In ihrer Studie achten sie darauf, welchen Teilen eines Tweets die meiste Aufmerksamkeit zukommt und wie die Aufmerksamkeitsallokation das Erinnerungsvermögen an diese Tweets beeinflusst (Counts & Fisher, 2011). Sie stellen fest, dass Nutzer nur eine sehr geringe Zeit aufwenden, um Tweets tatsächlich zu lesen und dass wenige Inhalte tatsächlich als relevant oder interessant gesehen werden und vieles auch sofort wieder vergessen wird (Counts & Fisher, 2011, S.7). Zudem zeigen sie, dass der Autor — Tweets von Freunden bleiben länger im Gedächtnis — starken Einfluss darauf hat, ob sich Nutzer an Tweets erinnern (Counts & Fisher, 2011, S.7).

Studien in diesem Kontext untersuchen auch, welche Charakteristika eines Tweets zu dessen themenbezogener/thematischer (*topical*) Relevanz beitragen (Hurlock & Wilson, 2011; Tao et al., 2012b). Zu inhaltlichen oder themenbezogenen Features sind aufgrund des Echtzeitcharakters von Twitter auch zeitbezogene Features Inhalt von Untersuchungen (Efron et al., 2014; Liang et al., 2014). Liang und Kollegen entwickeln ein Verfahren, das das Ranking eines Tweets erhöht, wenn dieser zeitlich gesehen nahe an dem *burst* eines bestimmten Themas ist (Liang et al., 2014).

Natürlich spielen auch twitterinterne Features eine Rolle, z. B. die Häufigkeit, wie oft ein Tweet weitergeleitet wurde oder ob Hashtags und URLs im Tweet vorhanden sind (Duan et al., 2010; Efron & Winget, 2010). Duan und Kollegen verwenden die Präsenz einer URL als Feature in einem Learning-to-Rank-Experiment (Duan et al., 2010). Efron und Kollegen verwenden in ihren Retrievalexperimenten das *relevance feedback* basierend auf Suchanfrageerweiterungen mit Hashtags (Efron, 2010). Neuste Verfahren zum Ranking von Tweets in der Twitter-Suche nutzen Verfahren wie *word embedding* (X. Yang et al., 2016) oder *deep learning* (Severyn & Moschitti, 2015). Abschließend bleibt zu erwähnen, dass nicht nur die Systemperspektive analysiert wird. Auch die Nutzer-Perspektive erhält Aufmerksamkeit. Dies geschieht in Form von eigenen Tweet-Suchmaschinen (Tao et al., 2012a) oder SUI-Erweiterungen wie Facetten-Suche für Tweets (Abel et al., 2011).

Themen, die der Microblogsuche nahestehen, wie die automatische Generierung einer Zusammenfassung von Twitter-Inhalten (*microblog summarization*) (Sharifi et al., 2010; Nichols et al., 2012; W. Xu et al., 2013; Shou et al., 2013), Verfahren, die über das Ranking von einzelnen Tweets hinausgehen und ganze Konversationen empfehlen (J. Chen et al., 2011) sowie die Vorhersage von RTs bzw. das Re-ranking der eigenen *Timeline*, sollen nicht unerwähnt bleiben, werden aber

bis auf zuletzt genanntes Thema, welches in Abschnitt 3.2.1.3 noch näher beleuchtet wird, aus Platzgründen nicht näher ausgeführt.

Eine ganzheitliche Diskussion der Verfahren, die für Retrievalexperimente verwendet werden, kann in dieser Arbeit nicht geleistet werden. Dennoch sind die angesprochenen Arbeiten zumindest in ihrer experimentellen Herangehensweise und der Nutzung bzw. Identifikation von Features von Interesse, da eine Twitter-Suche auch darauf ausgerichtet sein könnte, Re-finding zu unterstützen, eine Perspektive, die in Abschnitt 3.2.1.3 ausführlicher dargestellt wird und sich auf die hier präsentierten Arbeiten stützt.

### 3.2.1.2. Social Media als soziale Frage-Antwort-Systeme

Eine weitere Perspektive des Informationssuchverhaltens bzw. des Auffindens von relevanten Informationen lässt sich in der Nutzung von Social-Media-Applikationen als soziale Frage-Antwort-Systeme feststellen. Nutzer verleihen ihrem Informationsbedürfnis nicht Ausdruck, indem sie eine Suchanfrage formulieren, sondern indem sie eine Frage an ihre Freunde bzw. Followees senden. Ähnlich wie Microblogsuche im Allgemeinen wurde diese Variante des *social search* mit der Motivation beforscht, klassische Suchmaschinen-Ergebnisse noch stärker kontextualisieren und personalisieren zu können. Denn häufig erwarten Nutzer, die Fragen stellen, explizit Empfehlungen und Meinungen und weniger Fakteninformationen; ein Gebiet auf dem Suchmaschinen ohnehin bereits zufriedenstellende Ergebnisse liefern (Ringel Morris et al., 2010). Die Fragestellungen bei der Analyse des *status message question asking* (SMQA) sind von starker Heterogenität geprägt und reichen von der generellen Beschreibung dieses Suchverhaltens (Frequenz des Vorkommens, Motivation etc.) über die inhaltliche Bandbreite der gestellten Fragen, bis hin zu linguistischen Besonderheiten der formulierten Fragen, wie twitterspezifischen Sprachelementen oder dem Frageverhalten von bestimmten Nutzergruppen wie Journalisten. Schließlich werden auch die Qualität der Antworten und deren Relevanz für den Fragenden analysiert.

Zu den ersten Studien, die *social question answering* auf Social-Media-Plattformen analysieren, zählen die Studien von Efron und Winget sowie Ringel Morris, Teevan und Panovich (Efron & Winget, 2010; Ringel Morris et al., 2010). Efron und Winget erarbeiten auf der Basis von zwei Tweet-Korpora, die sie mithilfe der Twitter Streaming API erstellen, eine erste Taxonomie von Fragearten, die auf Twitter gestellt werden. Die gesammelten Frage-Tweets werden von Experten in vorgegebene Fragekategorien klassifiziert (Efron & Winget, 2010, S.5). Schließlich stellen sie eine zweidimensionale Schematisierung entlang der Dimensionen Informationsbedürfnis (sofort vs. beständig) und Leserkreis (Individuen vs. Netzwerk) (Efron & Winget, 2010). In einer Cross-Plattform-Studie (Twitter, Facebook) analysieren Ringel Morris, Teevan und Panovich neben Themenkategorien, zu denen Fragen gestellt werden, auch das Antwortverhalten der Followees: die Qualität der Antworten und die Gründe von Nutzern, weshalb sie selber auf solche Fragen antworten (Ringel Morris et al., 2010). Die Probanden ihrer Fragebogenumfrage geben zu etwa 70% an, bereits einmal eine relevante Antwort auf eine auf Twitter oder Facebook gestellte Frage erhalten zu haben. Bei der Analyse der Wahrnehmung der Antwortschnelligkeit geben Nutzer an, sehr schnelle Antworten zu erwarten, diese aber oft erst innerhalb eines Tages eintreffen. Dies liegt aber laut Angabe der Probanden noch in einem zeitlich vertretbaren Rahmen (Ringel Morris et al., 2010, S.1745). Paul, Hong und Chi nutzen die Crowdsourcing-Plattform *Amazon Mechanical Turk* (MTurk), um Fragen in via Twitter-API gesammelten Tweets zu identifizieren (Paul et al., 2011). Wie in anderen Studien stehen die Art der Fragen, deren inhaltliche Bandbreite sowie die Antwortanalyse im Fokus ihrer Untersuchungen. Final nutzen sie ein logistisches Regressionsmodell, um Faktoren zu eruieren, die die Antwortchancen erhöhen, wobei die Anzahl der Follower eines Nutzers sowie dessen Anmeldezeitpunkt auf Twitter gute Features bei der Vorhersage darstellen

(Paul et al., 2011). Yang et al. untersuchen die kulturellen Verschiedenheiten bei der Nutzung von Social-Media-Applikationen für *social search* unter Verwendung einer Fragebogenstudie (J. Yang et al., 2011). Asiatische Teilnehmer der Studie scheinen dieser Art des Informationssuchverhaltens aufgeschlossener gegenüber zu stehen als westliche Teilnehmer (J. Yang et al., 2011). Liu und Jansen vergleichen Fragen, die an klassische Suchmaschinen gerichtet werden mit Fragen, die an das Twitter-Netzwerk gesendet werden. Es zeigen sich deutliche Unterschiede in den Themen, die von den jeweiligen Kulturkreisen präferiert werden (Z. Liu & Jansen, 2012). Oeldorf-Hirsch et al. untersuchen experimentelle Situationen, in denen Nutzer zwischen einer Suchanfrage an eine normale Suchmaschine oder einer Frage an eine Social-Media-Plattform entscheiden sollen (Oeldorf-Hirsch et al., 2014). Das Laborexperiment zeigt, dass nur in 7% der Fälle Nutzer ausschließlich den Freundes-/Follower-Kreis einer Social-Media-Plattform befragen würden (Oeldorf-Hirsch et al., 2014). Bei der Analyse der Motive für die Nutzung in diesen Situationen stellen sich jedoch die selben Ergebnisse fest, wie in vergleichbaren Studien: stärkeres Vertrauen, bessere Kontextualisierung und Personalisierung durch den persönlichen Bezug zu den Followern sowie der Glaube, dass Suchmaschinen in diesen Situationen nicht funktionieren würden (Oeldorf-Hirsch et al., 2014, S.19). Hasanain, Elsayed und Magdy identifizieren Fragen in einer Stichprobe aus arabischen Tweets in einem Zwei-Stufen-Prozess: einem regelbasierten Filter sowie einem binären Klassifikator (Hasanain et al., 2014). Aktuelle Untersuchungen im Kontext *social media question answering* konzentrieren sich auf diese Praxis bei bestimmten Personengruppen wie Journalisten (Hasanain et al., 2016) oder vergleichen SMQA sogar mit dem Rat realer Personen (Fuchs & Groh, 2015).

Im Kontext der Untersuchung des Wiederfindensverhaltens spielt das Informationsverhalten des *social media question answering* wohl eine untergeordnete Rolle, wobei Nutzer aber auch Informationsbedürfnisse des Wiederfindens via Frage formulieren könnten.

### 3.2.1.3. Informationsdiffusion, RT-Vorhersage, Tweet-(Re)-Ranking und *Serendipity*

Das Auffinden von Informationen auf Social-Media-Plattformen steht in engem Zusammenhang mit der dortigen Informationsdiffusion, also der Verbreitung der Information über das Netzwerk der Plattform. Auf Twitter kann die Retweet-Funktion dazu genutzt werden, um Tweets weiterzuleiten und somit zur Verbreitung/Diffusion von Information beitragen. Im Folgenden werden Arbeiten besprochen, die mit Informationsdiffusion in Beziehung stehen, dies sind:

1. Arbeiten, die sich generell mit der Modellierung der Informationsdiffusion auseinandersetzen und sich dabei auch Modellen der Virologie bedienen.
2. Arbeiten, die Faktoren von weitergeleiteten Tweets analysieren und prädiktive Modelle erstellen, welche erklären, weshalb bestimmte Tweets stärkere Verbreitung finden als andere und die Verbreitung von Tweets vorhersagen.
3. In engem Zusammenhang mit den Faktoren die RTs beschreiben, steht das Re-Ranking von Tweets bzw. das Filtern der eigenen *Timeline*, da hier die Tatsache, dass ein Nutzer einen Tweet weiterleiten würde, als Grundlage genutzt wird und ein Retweet als Proxy für Interesse seitens des Nutzers gewertet wird.
4. Es steht die Frage nach der durch Informationsdiffusion zufällig entdeckten Information in Form von *Serendipity*-Erfahrungen der Nutzer im Raum.

Hoang führt in seiner Dissertation 44 Arbeiten an, die die Informationsdiffusion auf Twitter auf empirische Weise untersuchen oder modellieren (Hoang, 2016). Für den Kontext der hier behandelten PIM-Fragestellungen steht bei der Thematik Informationsdiffusion vor allem der globale Lebenszyklus von Tweets und deren Informationen im Vordergrund. Es wird also die Frage thema-

tisiert, wie lange sich Informationen auf Twitter im Umlauf halten. Dies ist zum einen interessant, da man vermuten würde, dass je länger Informationen präsent bleiben, desto eher sind sie für ein Weiterleiten auch wieder gefunden worden. Zum anderen steht dem globalen Lebenszyklus eines Tweets der eher individuelle, nutzerbezogene Lebenszyklus eines Tweets für Individuen gegenüber, der in dieser Arbeit verstärkt thematisiert wird (vgl. Abschnitt 6.4.4.3).

Bei der Analyse der Frage, ob Twitter ein *social network* oder ein Nachrichtenmedium ist, analysieren Kwak et al. zeitliche Muster von Retweets (Kwak et al., 2010). Sie erkennen, dass obwohl die Hälfte aller Weiterleitungen innerhalb von 24 Stunden geschehen, noch 10% der RTs nach einem Monat seit der Erstellung des Tweets passieren (Kwak et al., 2010, S.599). Wu et al. analysieren die Verbreitung von Information und deren Konsum auf Twitter, indem sie unter Verwendung der Twitter Listenfunktion Elite-Nutzer (Berühmtheiten, Blogger, Repräsentanten von Medien und Organisationen) mit normalen Nutzern vergleichen (Wu et al., 2011). Im Zuge ihrer Forschung untersuchen sie die Lebensdauer von Inhalten, welche als das erste und letzte Auftreten einer URL in einem auf der Plattform auffindbaren Tweet definiert ist. Sie stellen fest, dass einige Inhalte eine sehr kurze Lebensdauer besitzen und die Wahrscheinlichkeit, dass diese erneut in Twitter auftauchen, eher gering ist (Wu et al., 2011, S.713). Auf der anderen Seite des Kontinuums identifizieren sie Musikvideos oder Artikel aus Zeitschriften, die von normalen Nutzern gepostet werden, als Artefakte, die eine beinahe unbegrenzte Lebensdauer haben können (Wu et al., 2011, S.713). Sie argumentieren, dass die Langlebigkeit dieser Inhalte nicht durch Informationsdiffusion (RT) erklärt werden kann, sondern auf einem erneuten Entdecken dieser Inhalte abseits der Plattform beruhen. Bei Inhalten, die von Elite-Nutzern stammen, kann jedoch das Gegenteil beobachtet werden. Die Inhalte leben vielfach durch Informationsdiffusion weiter (Wu et al., 2011, S.713). Dies würde auch für ein erneutes Wiederfinden dieser Inhalte sprechen. Es lässt sich erkennen, dass sowohl Inhalt als auch Quelle Einfluss auf den Lebenszyklus von Information auf Twitter besitzen (Wu et al., 2011). Die Studien von Kwak und Kollegen sowie Wu et al. zeigen, dass Tweets eine durchaus lange Lebensdauer besitzen, was gegen die Kurzlebigkeit spricht, die dem Medium oft zugesprochen wird.

Betrachtet man die Lebensdauer von Inhalten auf anderen Social-Media-Plattformen, zeigen Chang et al., dass 95,8% aller Re-blogs auf Tumblr innerhalb von 24 Stunden stattfinden, und schreiben der Plattform einen starken Fokus auf *recency* zu (Chang et al., 2014, S.27). Dahingegen argumentieren Xu et al., dass die durchschnittliche Lebensdauer eines Tumblr Posts deutlich länger ist, als auf vergleichbaren Plattformen wie Twitter (J. Xu et al., 2014). Etwa 2,5% aller Posts haben eine Lebensdauer von über 40 Tagen (J. Xu et al., 2014, S.16). Großangelegte Twitter-API-Studien zeigen, dass die Wahrscheinlichkeit der Weiterleitung eines Tweets mit der Zeit seit seiner initialen Erscheinung in der *Timeline* des Nutzers sinkt (Galuba et al., 2010; Hodas & Lerman, 2012). Hodas und Lerman demonstrieren, dass das schnelle Verschwinden von Tweets aus der *Timeline* eines der Haupthindernisse für eine Informationsverbreitung im Netzwerk ist (Hodas & Lerman, 2012). Die Tatsache, dass Tweets schnell die Aufmerksamkeitsspanne des Nutzers verlassen, stellt eine natürliche Grenze für die Lebensdauer von Tweets dar und erschwert die Informationsverarbeitung im Allgemeinen. Ihr Argument stützen sie durch Analysen, die zeigen, dass das absolute Alter eines Tweets oder seine Interessanztheit keinen Einfluss auf die RT-Wahrscheinlichkeit besitzen. Yang und Counts unterstützen dieses Argument ebenfalls, indem sie zeigen, dass der allererste Tweet zu einem Thema nicht der Wichtigste ist, um weitere Abkömmlinge bzw. Informationsverbreitung zu erzeugen (J. Yang & Counts, 2010). Fasst man die letzten Abschnitte zusammen, gibt es einige Faktoren, die dafür sprechen könnten, dass ältere Inhalte wiedergefunden werden können und erneut virale Verbreitung finden.

Neben zeitlichen Faktoren untersuchen Forscher noch weitere Features, die bei der Vorhersage eines RT eingesetzt werden. Wie bereits weiter oben erwähnt, werden die Studien auch im Kontext

Timeline-Re-Ranking oder -Filtering durchgeführt, wobei die Annahme besteht, dass Tweets, die von einem Nutzer weitergeleitet werden, weiter oben in einer solchen neuen *Timeline* stehen sollten. Zu den Features zählen: Textinhalte, Tweet-Metadaten, Eigenschaften des Autors, wie eine hohe Followerzahl, die Häufigkeit in Twitter-Listen geführt zu sein oder die Anzahl der Favorisierungen des Tweets (Suh et al., 2010; Hong et al., 2011; Petrovic, Osborne, McCreadie et al., 2013; Tan et al., 2014). Während diese Untersuchungen hauptsächlich eine Vorhersage für eine breite Masse an Twitter-Nutzern durchführen, ist für einen so derart personalisierten Kontext wie PIM die Frage interessanter, ob ein Individuum einen Tweet weiterleiten würde. Auch hierzu werden Studien durchgeführt (Macskassy & Michelson, 2011; Uysal & Croft, 2011; Feng & Wang, 2013).

*Serendipity*, also die zufällige Begegnung mit relevanter Information, während man auf der Suche nach Information zu einem anderen Thema ist, ist ein entscheidender Bestandteil Twitters, das, so Mitgründer Jack Dorsey, explizit auf diese Situationen ausgerichtet ist: „How do we introduce you to something that would otherwise be difficult for you to find, but something that you probably have a deep interest in discovering?“ (Carmody, 2011). Studien zu *serendipity* sind methodisch eher qualitativ orientiert, wobei Studienteilnehmer in Interviews oder Fragebögen Situationen, in denen sie zufällig relevante Information finden, explizit kommunizieren. Beispiele hierfür sind die Studien von Dantonio, Makri und Blandford oder Bogers und Björneborn (Dantonio et al., 2012; Bogers & Björneborn, 2013). Dantonio et al. stellen in den geführten Interviews fest, dass es häufiger zu Serendipitysituationen auf Twitter kommen kann, wobei die Häufigkeit dieser Situationen mit der Nutzungsdauer korreliert (Dantonio et al., 2012). Bogers und Björneborn nehmen eine Inhaltsanalyse von Tweets vor, die mit dem Hashtag *#serendipity* getaggt sind. Eine großangelegte quantitative API-Studie zu unerwarteter Relevanz in Retweets führen Sun, Ming und Mei durch (Sun et al., 2013). Ihren Analysen nach enthalten 27% aller RTs für den Rezipienten unerwartete Information. Sie stellen fest, dass das Entdecken von Tweets mit unerwarteter Relevanz zu mehr Nutzeraktivität und sozialer Interaktion führt.

Die RT-Funktion hat zumindest implizit eine starke Bedeutung für die Abschätzung, wie häufig Wiederfinden auf Twitter stattfindet. Tatsächlich zeigt aber die Studie von boyd et al., dass die Funktion auch als Aufbewahrungsinstrument genutzt wird, was Abschnitt 3.2.2.2 zeigen wird.

#### 3.2.1.4. Informationserschließung und Hashtagnutzung

In Abschnitt 1.3, bei der Vorstellung von Twitter und seinen Funktionalitäten, wurden Hashtags als vom Nutzer erstellte Indexterme zur Informationserschließung charakterisiert. Tatsächlich erfüllen diese Indexterme nicht nur den Zweck der Informationserschließung. Hashtags dienen auch der Informationsdiffusion und Dialogförderung. Studien zur Hashtagverwendung sind ähnlich heterogen angelegt wie Studien zu Twitter selbst. Sie untersuchen die Nutzung des Hashtags von bestimmten Personengruppen, wie Politikern oder Journalisten (Small, 2011) oder während spezifischen Ereignissen, wie Krisensituationen bzw. zur Ereigniserkennung im Allgemeinen (Cui et al., 2012). Ein sehr beliebtes Themengebiet sind Studien, die prädiktive Modelle erstellen, um die Beliebtheit und Verbreitung von Hashtags vorherzusagen (L. Yang et al., 2012; Ma et al., 2013; Pervin et al., 2015) und hieraus schließlich auch Empfehlungen für den Nutzer zu erstellen (Otsuka et al., 2014). Von Interesse ist auch der dialogorientierte Charakter, den Tweets durch Hashtags erhalten — Teilhabe an einer gemeinschaftlichen Diskussion durch Verwendung eines Hashtags — und den Beitrag, den Hashtags zur Verbreitung eines Tweet-Inhalts liefern (J. Huang et al., 2010), die Kookkurrenz von Hashtags (Pervin et al., 2015) und ob diese häufig wieder verwendet werden (Ma et al., 2012). Das Wiederfinden eines Tweets, um die korrekte Schreibweise eines Hashtags zu ermitteln, könnte bei einer Wiederverwertung ein sinnvolles Nutzerverhalten darstellen. Auch die linguistische Perspektive hinter Hashtags, wie das Tagging-Vokabular, wird in Forschungsarbeiten

thematisiert (Cunha et al., 2011; Zappavigna, 2015).

Huang, Thornton und Efthimiadis vergleichen die Nutzung von Hashtags auf einer klassischen Social-Tagging-Plattform (*Delicious*) mit der Hashtag-Praxis auf Twitter (J. Huang et al., 2010). Sie stellen fest, dass auf Twitter ein neues Tagging-Phänomen entsteht, dass sie als *micro-meme* bezeichnen, wobei die Verwendung eines Tags *a-priori* erfolgt und nicht wie beim klassischen Tagging *a-posteriori* (J. Huang et al., 2010). Aus ihrer Perspektive werden Hashtags auf Twitter hauptsächlich zur Teilhabe an öffentlichen Diskussionen genutzt und nicht, um Inhalte zu organisieren oder abzufragen (J. Huang et al., 2010).

Romero, Meeder und Kleinberg analysieren, ob und wie die Art und Weise der Hashtag-Verwendung eines Individuums auch die Nutzung dieses Hashtags im Follower-Followee-Netzwerk beeinflusst (Romero et al., 2011). Unterschiedliche Themen und Arten von Hashtags verbreiten sich unterschiedlich schnell bzw. bleiben länger im Umkreis als andere. Beispielsweise haben politische Hashtags eine längere Halbwertszeit als Hashtags anderer Themen (Romero et al., 2011, S.703). Dies wird auch durch eine Studie von Lin et al. bestätigt (Y.-R. Lin et al., 2013).

Yang et al. schreiben dem Hashtag eine Bookmarking- und Organisationsfunktion zu, charakterisieren ihn aber gleichzeitig auch als gemeinschaftsbildendes Objekt, so dass die Verwendung eines Hashtags sowohl Interessensbekundung zu einem Thema, aber auch Zugehörigkeit zu einer Community ausdrückt (L. Yang et al., 2012). Mithilfe von maschinellen Lernverfahren versuchen sie, zukünftige Hashtag-Nutzung vorherzusagen und somit den Weg für ein Hashtag-Empfehlungssystem zu bereiten (L. Yang et al., 2012). Otsuka, Wallace und Chiu führen ein neues Ranking-Verfahren (*hashtag frequency-inverse hashtag ubiquity* (HF-IHU)) ein, welches die Relevanz eines Hashtags und die Spärlichkeit der gegebenen Daten in Betracht zieht, um Hashtag-Empfehlungen zu liefern (Otsuka et al., 2014). Posch et al. analysieren, inwieweit die pragmatischen Charakteristika von Hashtags etwas über deren Semantik aussagen (Posch et al., 2013). Im Kontext von Marketing-Studien wird erforscht, inwieweit Tweets mit Hashtags mehr Klicks erreichen. Interessanterweise stellt Twitter bei eigenen Studien fest, dass Tweets ohne Hashtags und *@Mentions* 23% mehr Klicks erhalten (Wagner, 2015).

Hashtags finden nicht nur auf Twitter Verwendung. Auch auf anderen Social-Media-Plattformen werden vergleichbare Studien zur Hashtag-Verwendung durchgeführt. Ferrara, Interdonato und Tagarelli untersuchen — neben anderen Charakteristika der Plattform — das Tagging-Verhalten der Nutzer auf Instagram (Ferrara et al., 2014). Booten analysiert das Phänomen des *hashtag drift*, also die Zunahme an zusätzlichen weiteren Hashtags zu einem zentralen Hashtag im Kontext politischer Mobilisierungen auf Tumblr (Booten, 2016).

Die letzten Abschnitte haben gezeigt, dass, obwohl Tags respektive Hashtags, die in ihrem ursprünglichen Gebrauch dazu gedacht sind, PIM zu betreiben und die Organisation und das Auffinden (potentiell auch das erneute Auffinden) von Informationen zu erleichtern, auf Twitter eine gänzlich andere Funktion einnehmen. Immer wieder wird in den Studien der gemeinschaftsbildende Charakter von Hashtags und die dadurch entstehenden Konversationen betont. Dies trägt jedoch kaum zu einem leichteren Management von Tweets bei. Offensichtlich ist aufgrund der Masse an Tweets, die identische Hashtags enthalten, ein Management/Wiederfinden via Hashtags kaum möglich. Jedoch wird im Kontext von Re-finding — in Abschnitt 3.2.3 — erwähnt werden, dass das Vorhandensein eines Hashtags das Wiederfinden erleichtern kann, was qualitative Studien zeigen (Elsweiler & Harvey, 2015).

### 3.2.2. Social-Media-Nachrichten Verwalten

In diesem Abschnitt werden Forschungsarbeiten vorgestellt, die in Beziehung zu Managementverhalten auf Social-Media-Plattformen stehen. Hierzu gehören Verhaltensweisen, wie das Zusammenstellen von Social-Media-Inhalten zu Kollektionen sowie deren Organisation und das Aufbewahren (*keeping*) von Social-Media-Posts im Allgemeinen.

Eine Perspektive, die PIM und Social Media verbindet, betrifft das Management des Selbstbildes, sowohl im Hinblick auf die Eigenwahrnehmung als auch auf die Fremdwahrnehmung. In der Forschung wird dies als *self presentation*, *impression management*, oder *reputation management* bezeichnet und kann in den breiteren Kontext von Social-Media-Nutzung und Privatsphäre (*privacy*) eingebettet werden (Hogan, 2010; Y. Wang et al., 2011; Litt et al., 2014; Litt & Hargittai, 2014; Woodruff, 2014). Charakteristika des Leserkreises wie Größe oder Verschiedenheit (*imagined audience*) sowie Langlebigkeit und Zugangsmöglichkeiten von Inhalten stellen Faktoren dar, die das Management des Selbstbildes auf Social-Media-Plattformen komplex und schwierig gestalten (Marwick & boyd, 2011; Litt & Hargittai, 2014). Tatsächlich hat ein Nutzer von Social-Media-Anwendungen die Gestaltung seines Selbstbildes nicht komplett in eigener Hand. Alles was zum Selbstbild beiträgt, wird auch von anderen Nutzern und deren Interaktionen mit dem eigenen Account bzw. dessen Inhalten geprägt (Walther et al., 2009; Litt et al., 2014). Für den PIM-Kontext und der Verwaltung des Selbstbildes stehen jedoch vor allem die eigenen Handlungen und insbesondere die eigenen gesendeten Nachrichten, Posts oder Einträge und der Umgang mit ihnen im Vordergrund.

Eine ganz konkrete Form des *impression management* stellt die sog. *regret deletion* dar. *regret in social media* bzw. *regret deletion*, bezeichnet das Verhalten von Nutzern, die bedauern, bestimmte Nachrichten gepostet zu haben, und auf dieses Bedauern entsprechend reagieren, z. B. indem sie diese Nachrichten löschen (Y. Wang et al., 2011; Sleeper, Cranshaw et al., 2013; Litt et al., 2014; Almuhimedi et al., 2013; Zhou et al., 2016). Ziel dieses Managements sind beispielsweise Nachrichten, bei denen die Autorin ihre Meinung ändert, oder Nachrichten, die private oder zu emotionale Informationen enthalten bzw. beleidigend sind gegenüber Personen oder Gruppen. Als Beispiel können Partyfotos angeführt werden, die aus einer freudigen Laune heraus gesendet wurden, im Nachhinein aber als zu unpassend identifiziert werden.

Die Analyse von Gründen, die das Bedauern auslösen und entsprechende Nutzerreaktionen hervorrufen, finden sich für Facebook bei Wang und Kollegen sowie für Twitter in einer Arbeit von Sleeper et al. (Y. Wang et al., 2011; Sleeper, Cranshaw et al., 2013). Hervorzuheben ist dabei, dass der Impuls für das Management dieser Tweets nicht nur intrinsisch veranlasst sein muss, sondern der Auslöser auch von außen, durch einen verbalen Hinweis oder eine Tat von anderen, kommen kann. Es gibt eine Reihe an sog. *repair strategies*, um auf dieses Bedauern entsprechend zu reagieren, wie eine Entschuldigung oder eine Rechtfertigung (Sleeper, Cranshaw et al., 2013, 3283). Als Prävention, damit solche Situationen erst gar nicht auftreten, halten sich Nutzer ständig dazu an, Inhalte zu bewerten und abzuwägen, welche gepostet oder weitergeleitet werden und welche nicht. Ein Verhalten, das auch als *self-censorship* oder *information disclosure* bekannt ist (Das & Kramer, 2013; Sleeper, Balebako et al., 2013). Systemseitige Unterstützung könnte Nutzer vor dem Senden derartiger Tweets warnen und eine spätere Reuesituation vermeiden (Zhou et al., 2016). Aufgrund des starken persönlichen und subjektiven Charakters derartiger Tweets, stellt die Identifikation von möglichen Regret-deletion-Tweets ein komplexes Problem dar (Zhou et al., 2016). Hierbei ist jedoch auch eine Brücke zum Wiederfindensverhalten zu vermerken, da wiederfindenswerte Tweets vermutlich eine ähnlich starke persönliche und subjektive Färbung besitzen.

Das Löschen einer Nachricht stellt dasjenige Verhalten dar, das dem Verwalten und auch Wiederfinden von Tweets am nächsten kommt, da, um einen Tweet löschen zu können, dieser zuerst

lokalisiert werden muss. Hierbei muss erwähnt werden, dass das Löschen eines Tweets aber nicht ausschließlich durch den Nutzer initialisiert werden muss, da auch Twitter im Rahmen ihrer *terms of service* inadäquate Tweets entfernt. Zudem können mehrere Varianten des Löschens existieren. Beispielsweise werden, wenn ein Account gelöscht wird (durch Nutzer oder Twitter), alle Tweets des Accounts entfernt (Petrovic, Osborne & Lavrenko, 2013, S.1). Die Dimension des *last-minute self-censorship*, also das Ausformulieren von Nachrichten, die dann aber im letzten Moment zurückgehalten werden, also nicht gesendet werden, wird nicht betrachtet (Das & Kramer, 2013).

Almuhimedi et al. untersuchen in einer großangelegten, quantitativen Studie das Löschverhalten von Nutzern und die Charakteristika von 1,6 Millionen gelöschten Tweets (Almuhimedi et al., 2013). Zum Zeitpunkt der Studie war es Drittanbietern noch möglich, Archive von Tweets anzulegen, was seit einer Änderung der Twitter *terms of service*<sup>2</sup> (AGBs) nicht mehr erlaubt ist und auch dazu geführt hat, dass zahlreiche Dienste, die eine Archivierung von Tweets vorgenommen haben, geschlossen wurden (Fingas, 2016). Almuhimedi und Kollegen stellen fest, dass die Hälfte aller Nutzer Tweets löscht. Besonders häufig sind einfache Status-Updates und Tweets, die eine @Mention enthalten, von Löschungen betroffen. Bei Tweets, die eine Antwort (@Reply) erhalten, sehen Nutzer eher davon ab, diesen zu löschen (Almuhimedi et al., 2013, S.901). Auf inhaltlicher Ebene unterscheiden sie zwischen oberflächlichen (*superficial*) Löschungen und der bereits oben erwähnten Klasse der *regret deletions*. Einfache Löschungen sind dadurch charakterisiert, dass Nutzer Rechtschreibfehler korrigieren wollen oder fehlende @Mentions oder Links hinzufügen, wobei sie den initialen Tweet löschen und eine verbesserte Version des Tweets senden (Almuhimedi et al., 2013, S.901). Sie untersuchen vier stereotypische Kategorien von bedauerten Tweets, welche mithilfe eines lexikalischen Ansatzes kategorisiert werden. Jedoch können sie sowohl in Bezug auf die Anzahl der aus diesen Kategorien stammenden Tweets als auch in Bezug auf das Sentiment der Kategorien keine Unterschiede zwischen gelöschten und ungelöschten Tweets ausgemacht werden (Almuhimedi et al., 2013, S.902). Eine zeitliche Analyse zeigte, dass im Durchschnitt Tweets nach 8,5 Std. gelöscht werden, jedoch der Wochentag keinen Einfluss auf das Löschverhalten der Nutzer hat (Almuhimedi et al., 2013, S.903).

Eine weitere Arbeit, die das Thema Löschverhalten thematisiert, ist die Arbeit von Petrovic, Osborne und Lavrenko, die ebenfalls auf eine Vorhersage von gelöschten Tweets abzielen und dabei soziale Features (Account-Informationen wie Anzahl Follower, Anzahl Following, Anzahl der Tweets sowie Anzahl Hashtags oder Mentions), Autoren-Features und Text-Features (alle Wörter in einem Tweet) verwenden (Petrovic, Osborne & Lavrenko, 2013). Xu et al. konzentrieren sich im Kontext *regret deletion* auf eine ganz besondere Klasse von Tweets, nämlich *bullying Tweets*, also Tweets, die andere Personen schikanieren oder einschüchtern und dann später aus Reue gelöscht werden (J.-M. Xu et al., 2013). Bagdouri und Oard führen ein ähnliches Vorhersageexperiment durch wie Petrovic et al. bzw. Xu et al., um mit dem Wissen, welche Tweets gelöscht werden, eine lückenlose Testkollektion für zukünftige Experimente zu erstellen (Bagdouri & Oard, 2015).

Die bis dato aktuellste Untersuchung zum Löschverhalten stellt die Arbeit von Zhou, Wang und Chen dar, welche gezielt die Charakteristika und Identifikation von *regret deletions* via Merkmale des Tweet-Inhalts (*content-indentifiable regrettable Tweet*) in den Fokus stellt (Zhou et al., 2016). Um Löschverhalten zu verstehen, nehmen sie ein Clustering von Nutzeraccounts vor, die durch fünfdimensionale Feature-Vektoren repräsentiert sind, um — ihrer Annahme nach — einer gewöhnlichen Nutzergruppe mit klassischen Regret-deletion-Verhalten am nächsten zu kommen (Zhou et al., 2016, S.605). Um Gründe für das Bedauern des Sendens zu analysieren und diese schließlich in einem Vorhersageexperiment nutzen zu können, verwenden sie das Kategorienschema von Wang et al. (Y. Wang et al., 2011), wenden dieses auf 4000 gelöschte Tweets an und klassifizieren die Gründe nochmals in zehn Kategorien. Im Vorhersageexperiment werden schließ-

<sup>2</sup><https://twitter.com/tos?lang=de>



lich inhaltsbezogene Features (Sentiment, Reuegrund basierend auf lexikalischen Features) sowie Nutzerpräferenzen (Sendeverhalten — Nutzer, die häufig Tweets mit Schimpfwörtern posten, tendieren dazu, diese nicht zu löschen — und Löschverhalten) als Features verwendet (Zhou et al., 2016, S.609). Es zeigt sich, dass die Verwendung von Nutzerpräferenzen als Features deutlich bessere Ergebnisse liefert als die Inhaltsfeatures alleine. Diese können jedoch nicht ganz vernachlässigt werden, weshalb Zhou und Kollegen sich in ihrer Annahme bestätigt sehen, dass unterschiedliche Nutzer, verschiedene Level des Bedauerns besitzen (Zhou et al., 2016, S.610).

### 3.2.2.1. Social Media als Archiv

In einer qualitativen Interview-Studie untersuchen Zhao und Kollegen aus welcher Motivation heraus und auf welche Art, Facebook-Nutzer ihre persönlichen Daten über die Zeit hinweg verwalten (Zhao et al., 2013). Ihre Analysen zeigen, dass im Umgang mit persönlichen Daten Facebook in drei unterschiedliche *Regionen* aufgeteilt werden kann, wobei *region* als ein Zusammenspiel aus verschiedenen Zielen, Bedürfnissen, Kontextfaktoren und den UI-Features der Plattform, die dieses Verhalten unterstützen, definiert wird (Zhao et al., 2013, S.1). Diese drei Regionen — inspiriert durch die Arbeit von Hogan (Hogan, 2010) — sind folgende:

- Die *performance region*: Bei ihr steht die Erstellung von Inhalten für einen spezifischen Echtzeit-Kontext mit aktueller Relevanz im Vordergrund (Zhao et al., 2013, S.4).
- Die *exhibition region*: Sie ist geprägt durch die Verwaltung des Selbstbildes („managing overall self-image“) über einen längeren Zeitraum hinweg (Zhao et al., 2013, S.5)). Besonders emotionale oder sensible Inhalte werden einem Check unterzogen und falls sie nicht den Ansprüchen entsprechen — da sie z. B. zu viel preisgeben, nicht mehr ihren Einstellungen entsprechen, oder zu beleidigend waren — werden diese möglicherweise anschließend gelöscht. Das Konzept der *regret deletion* wurde bereits in Abschnitt 3.2.2 beschrieben. Nicht nur emotionale Inhalte können das Ziel von Management-Aktivitäten sein. Nutzer versuchen generell die eigenen Inhalte (sowohl in Bezug auf Inhalt als auch auf Zeit) so relevant wie möglich zu halten.
- In der *personal region* steht der persönliche Wert von Daten für einen Selbst im Vordergrund. Diese Region hat am meisten Bedeutung für die Erstellung eines „personal locker“, also eines persönlichen Archivs via Facebook, das vor allem zum Schwelgen in Erinnerungen genutzt wird (Zhao et al., 2013). Auch hier motiviert die Verwaltung des eigenen Selbstbilds die Entscheidungen darüber, welche Inhalte wie lange im Archiv gehalten werden. Lebensverändernde Situationen, wie ein neuer Arbeitsplatz oder eine neue Beziehung, können Einfluss auf das Löschen bzw. Archivieren von Nachrichten haben.

Besonders bei den Regionen zwei und drei wird die Dimension Facebooks als persönliches Archiv evident. Ob nun in Bezug auf das Löschen von Nachrichten oder zum Schwelgen in Erinnerungen, es ist ein starker PIM-Bezug sowohl in Hinsicht auf das Management der Informationen als auch in Hinsicht auf Wiederfinden per se zu erkennen; denn bevor diese Informationen verwaltet werden können, müssen sie erst lokalisiert werden.

Zhao und Lindley fokussieren in einer weiteren Studie noch stärker auf die Perspektive von Social Media als persönliches Archiv (Zhao & Lindley, 2014). In einer Interviewstudie stellen sie fest, dass verschiedene Social-Media-Plattformen aus der Perspektive der Nutzer auch unterschiedlich stark als Archiv wahrgenommen werden. So charakterisieren Nutzer Facebook als diejenige Plattform, die als persönliches Archiv genutzt werden kann. Twitter bzw. Tweets hingegen hätten eine zu geringe Halbwertszeit (Zhao & Lindley, 2014, S.2434). Aufbauend auf dieser Beobachtung formulieren sie eine Nutzungsweise, welche sie als „*curation through use*“ bezeichnen. Dadurch, dass Nutzer

verschiedene Social-Media-Plattformen für unterschiedliche Zwecke nutzen und Inhalte auch gezielt je nach Wertigkeit verteilen — Facebook besitzt eine höhere Schwelle für eine neue Nachricht, da dort sehr häufig persönliche Daten gepostet werden, wohingegen Tweets auch eher belanglosen Inhalt enthalten können — entstehen unterschiedliche Nutzungsweisen, die auch unterschiedlich stark den Bezug zu persönlichen Archiven deutlich machen (Zhao & Lindley, 2014). Zhao und Lindley argumentieren außerdem, dass Nutzer aufgrund des aufgebauten Wertesystems für Social-Media-Applikationen ein gutes Gefühl dafür haben, wo bestimmte Informationen wiederzufinden sind (Zhao & Lindley, 2014, S.2438). Nichtsdestotrotz konnten sie feststellen, dass Social-Media-Applikationen, insbesondere Facebook, als Archive benutzt (verwalten von Nachrichten/Informationen) und interpretiert werden. Sie werden als Orte wahrgenommen, an denen Informationen nachhaltig für die zukünftige Nutzung gespeichert sind. Zusätzlich geben sie Designvorschläge und plädieren für eine Integration von online und offline Archiven für Nutzer.

Marshall und Shipman befassen sich ebenfalls mit der Frage nach Facebook als Archiv, um die Beständigkeit eines bedeutenden Teils der digitalen, persönlichen Daten einer Person sicherstellen zu können (Marshall & Shipman, 2014). In ihren Analysen steht jedoch weniger der Social-Media-Account als persönliches Archiv im Mittelpunkt, sondern die Frage, ob Facebook durch eine öffentliche Institution archiviert werden soll. Für das Szenario eines institutionellen Facebook-Archivs entwerfen sie hypothetische Situationen — nach der Methode *hypotheticals as heuristic device* von Rissland und Ashley (Rissland & Ashley, 1986) — für die Charakteristika dieses Archiv, um die Einstellungen bzw. Grenzen für Nutzer zu erheben (Marshall & Shipman, 2014, S.14).

Wiederum Marshall und Shipman analysieren in einer crowdsourcingbasierten Fragebogen-Umfrage Normen, Einstellungen und Grenzen der Nutzer gegenüber dem Speichern, Entfernen und der Wiederverwendung von Social-Media-Daten auf Facebook (Marshall & Shipman, 2015). Dabei steht weniger das konkrete Verhalten der Probanden im Fokus als deren Einstellung zu bestimmten hypothetischen Szenarien und Möglichkeiten beim Umgang mit Daten auf der Plattform. Beispielsweise versuchen sie zu ergründen, wie stark Nutzer der Situation zustimmen würden, dass bei der Erstellung einer Sicherungskopie des eigenen Accounts auch die Accounts (sämtliche Inhalte) der Freunde mitgespeichert werden (Marshall & Shipman, 2015, S.716), oder ob sie mit einem grundsätzlichen Weiterverkauf ihrer Daten an Drittunternehmen, wie Amazon, einverstanden wären (Marshall & Shipman, 2015, S.719). Sie stellen u.a. fest, dass soziale Distanz bzw. Nähe Einfluss darauf hat, ob Facebook-Daten gespeichert werden können sollten oder nicht.

Die Wahrnehmung von Social Media als persönliches Archiv, das auch wiederum zur Präsentation der eigenen Interessen für Außenstehende genutzt werden kann, wird vor allem *Pinterest*<sup>3</sup> zugeschrieben. Pinterest wird als Social-Collection- oder Social-Curation-Plattform wahrgenommen und von Probanden sogar als Lager oder Bibliothek für persönliche Kollektionen interpretiert (Hall & Zarro, 2012; Zarro & Hall, 2012; Han et al., 2014). In einer Studie von Zarro, Hall und Forte geben Nutzer an, es als PIM-Tool zu verwenden (Zarro et al., 2013). Bei Linder, Snodgras und Kerne wird es als Archiv bezeichnet (Linder et al., 2014). Nach Hall und Zarro verschwimmen auf Pinterest die Grenzen zwischen lokalen PIM-Kollektionen und öffentlichen Sammlungen, die auf Pinterest mit anderen geteilt werden (Hall & Zarro, 2012). Nutzer scheinen ihre persönlichen Favoriten auf Pinterest zu teilen, um diese anderen zu präsentieren bzw. eine Online-Kopie von ihnen zu erstellen (Hall & Zarro, 2012, S.7). Zarro, Hall und Forte führen eine qualitative Studie bestehend aus Interview und teilnehmender Beobachtung durch, um zu analysieren, wie Pinterest als Verwahrungsort genutzt wird (Zarro et al., 2013). Sie untersuchen das Nutzerverhalten sowie die Gründe für das Verhalten entlang der Dimensionen *collecting*, *discovering*, *collaborating*, *publishing*, *ownership* und *interaction and identity*. Ihre Probanden geben an, Sammlungen zu aktuellen Informationsbedürfnissen zu erstellen, sowie diese zu organisieren und zu verwalten. Sie beschreiben

<sup>3</sup><https://de.pinterest.com/>

Pinterest als visuelles Bookmarking-Werkzeug (Zarro et al., 2013, S.3). In ihren Analysen gehen Zarro und Kollegen jedoch nicht ins Detail in welchem Ausmaß, in welcher Frequenz und auf welche Weise Re-Organisation und Re-finding von Pins stattfindet. Linder, Snodgrass und Kerne charakterisieren Pinterest als Ideengenerierungsplattform bei der Nutzer Informationen für Kreativprozesse finden, aufbewahren und zu diesen zurückkehren (Linder et al., 2014). Beispielsweise führen sie die Aussage einer Probandin an, die zu gefundenen Rezepten zurückkehrt, um diese auszuprobieren und anschließend eine Re-kategorisierung vornimmt in eine Gruppe mit schmackhaften (Tried-it-loved-it-Seite) bzw. weniger schmackhaften (Tried-it-wouldn't-make-it-again-Seite) Rezepten (Linder et al., 2014, S.8). Rückkehr zu Pins kann unterschiedliche Motive besitzen, z. B. emotionale, um Motivation zu erhalten oder in Erinnerung zu schwelgen, oder praktische, um eine Liste wichtiger Dinge für einen Urlaub erneut zu verwenden (Linder et al., 2014, S.8).

Die Dimension *Social Media als Archiv* bezieht sich vor allem auf die von einem selbst erstellten, persönlichen Inhalte und das Management der eigenen Persönlichkeit. Tatsächlich stehen in dieser Perspektive ausschließlich die eigenen gesendeten Nachrichten im Fokus. Auf die Frage, wie Social-Media-Nachrichten von anderen in diese Kollektion integriert werden, ob diese auch aufbewahrt und wiedergefunden werden, bzw. ob und wie oft Nutzer das Bedürfnis haben, zu diesen zurückzukehren, wird nicht Bezug genommen.

### 3.2.2.2. Social-Media-Inhalte Aufbewahren

Erste Hinweise darauf, dass Nutzer das Bedürfnis haben, Tweets aufzubewahren, können in einer Analyse des RT-Verhaltens von boyd, Golder und Lotan ausgemacht werden (boyd et al., 2010). boyd sieht und analysiert die RT-Funktion vor allem als Möglichkeit, um einen größeren Personenkreis an Konversationen teilhaben zu lassen. Ihre Analysen des RT-Verhaltens führen sie anhand vier Datensätzen durch: (i) eine Sammlung von Tweets via API, (ii) eine Stichprobe von randomisiert ausgewählten Retweets, (iii) thematischen Geschichten bzw. Threads die aus Retweets bestehen gesammelt via Twitter-API und (iv) qualitative Kommentare aus Antworten auf Fragen, die via eines Twitter-Accounts an die Follower gesendet wurden. Neben der Frage *Wie* sich Nutzer beim Weiterleiten von Tweets verhalten — ein Verhalten, das sich im Laufe der Twitter-Evolution immer wieder geändert hat, weil sich auch die RT-Funktionalität verändert hat (vgl. Abschnitt 1.3) — werden die Fragen nach den inhaltlichen Schwerpunkten sowie den Gründen und Motiven für das Weiterleiten erforscht. Neben zahlreichen anderen Gründen, wie das Verbreiten eines Tweets an ein anderes Publikum, die Nutzung von RT als Kommentarfunktion oder um die Präsenz als Zuhörer zu signalisieren, ist der für den vorliegenden Kontext wichtigste Grund folgender: „To save tweets for future personal access (e.g. @peteaven: so I can find the tweet later by searching on myself, checking my updates)“ (boyd et al., 2010, S.6). Nutzer leiten Tweets weiter, um diese inmitten der eigenen gesendeten Tweets aufzubewahren und somit einen schnelleren Zugang zu diesen Tweets zu haben.

Gorrell und Bontcheva analysieren die Nutzung der Favoriten-Funktion, indem sie in einem Klassifikationsexperiment eine automatische Klassifikation von favorisierten Tweets in eine von fünf Kategorien zu erreichen versuchen (Gorrell & Bontcheva, 2016). Diese fünf Kategorien sind: Favorisieren als *Like*, zur Kommunikation von Dank, zur Selbstdarstellung, als Konversationsinstrument und zum Setzen von Bookmarks (Gorrell & Bontcheva, 2016, S.18). Interessant ist dabei, dass die Autorinnen der Nutzung von Favorisieren als Bookmark-Funktion nicht nur den Zweck des Aufbewahrens und der Erinnerungsfunktion zuschreiben, sondern das Bookmarking auch explizit als Möglichkeit der Interessenskommunikation via Favoritenliste interpretieren (Gorrell & Bontcheva, 2016, S.19). Für das Klassifikationsexperiment konstruieren sie zunächst einen Goldstandard aus 688 manuell klassifizierten Tweets. Anschließend vergleichen sie einen SVM-Klassifikator mit ei-

nem regelbasierten Baseline-System und der Variante, alle Favoriten als Likes zu interpretieren. Das Klassifikationsproblem wird fünf Mal als binäres Problem betrachtet. Als Features werden z. B. verwendet: Part-of-Speech-Charakteristika der Tweets, Existenz der *@Mention* des Favorisierenden im Tweet, RT-Häufigkeit des Tweets oder Präsenz einer URL. Der SVM-Klassifikator erreicht eine *accuracy* von 0,85 und ist deutlich besser als das regelbasierte System (*accuracy* = 0,77). Sie schließen mit den Aussagen, dass die Klassen *conversational* und *thanks* beinahe nicht zu unterscheiden sind, aber Nutzer beim Einsatz der Favoriten-Funktion hauptsächlich einen Anwendungszweck im Sinn haben und nicht zwischen unterschiedlichen Motivationen wechseln (Gorrell & Bontcheva, 2016, S.22). Diese Aussage wird in Abschnitt 5.6.3.4 nochmals aufgegriffen.

### 3.2.3. Social-Media-Inhalte Wiederfinden

Studien, die explizit das Wiederfindensverhalten von Nutzern auf Social-Media-Plattformen untersuchen, existieren nach bestem Wissen nicht. Forschungsarbeiten, welche zumindest indirekt das Thema des Wiederauffindens aufgreifen, wurden in den meisten Fällen bereits in Abschnitt 3.2.1.1 genannt. Dort wurde jedoch vor allem auf das Auffinden von Inhalten fokussiert. Im Folgenden stehen die Erkenntnisse der jeweiligen Forscher — wenn denn vorher schon aufgeführt — zum Thema Re-finding im Fokus.

Neben dem von boyd, Golder und Lotan festgestellten Verhalten, dass Nutzer Tweets weiterleiten, um diese später in den eigenen gesendeten Tweets leichter lokalisieren zu können, bringt die Studie *#TwitterSearch* von Teevan et al. weitere Hinweise, dass Nutzer Tweets wiederfinden (vgl. Abschnitt 3.2.1.1). Zwei Probanden der qualitativen Vorstudie geben an, via Suchanfragen zu vorher gesehenen Tweets zurückkehren zu wollen (Teevan et al., 2011, S.37). In der quantitativen Analyse der Suchanfragen stellen sie zwar fest, dass Suchanfragen sogar häufiger wiederholt werden als im Web-Kontext (55,76% aller Suchanfragen auf Twitter waren *repeated queries*), jedoch interpretieren sie diese Suchanfragen als *monitoring*-Verhalten, also als das Überwachen eines bestimmten Themas, indem die gleiche Suchanfrage immer wieder gestellt wird, um so die neuesten Tweets zu diesem Thema zu erhalten, anstelle von Wiederfinden (Teevan et al., 2011, S.40). In einem Lehrbuchartikel zur Erhebung und Analyse von Logdaten im HCI-Kontext greifen Dumais et al. nochmal die Suche nach bereits betrachteten Tweets via Twitter-Suche auf: „Similarly, people might want to search for old posts on Twitter, but this behavior is rarely observed [...]. This could be because old posts are not interesting, or it could be an artifact of the fact that the Twitter search interfaces currently only returns the most recent tweets.“ (Dumais et al., 2014, S.356). Sie vermuten, dass Wiederfinden deswegen so selten beobachtet wird, weil die Twitter-Suche das Verhalten nur sehr schlecht unterstützt, was jedoch nicht bedeutet, dass dieses Bedürfnis unter den Nutzern nicht vorhanden ist.

In Abschnitt 3.2.1.1 wurde bereits erwähnt, dass die Ergebnisse der Studie von Elsweiler und Harvey das Wiederfinden von Tweets als den Hauptmotivationsgrund für die Nutzung der Twitter-Suche hervorgebracht hat (Elsweiler & Harvey, 2015, S.274). 43% der Aufgaben, die über die Tagebuchstudie berichtet wurden, sind Wiederfindensaufgaben. In ihrem Kodierschema der Twitter-Suchaufgaben erfolgt eine weitere Differenzierung von Wiederfinden in die beiden Sub-Kodes *external resource linked to in tweet*, also Tweets, die Links, Bilder und Videos enthalten, und *content of tweet itself*, also Tweets bei denen der Tweet-Text im Vordergrund steht (Elsweiler & Harvey, 2015, S.269). Elsweiler und Harvey vertiefen die Diskussion um Re-finding-Verhalten weiter, indem sie zeigen, dass Re-finding eine schwierige Aufgabe ist, da Wiederfinden und Aufgabenmisserfolg in starkem Zusammenhang stehen. Nach Angaben ihrer Nutzer helfen Hashtags dabei, Wiederfindensaufgaben weniger komplex zu gestalten.

In einer Mixed-Methods-Studie zur partizipativen Hashtag-Verwendung am Beispiel des Weekend-

Hashtag-Projekts (*#WHP*) auf Instagram beobachten Oh und Kollegen das Phänomen des *Re-tagging*, welches die Rückkehr zu älteren, bereits geposteten Bildern beschreibt, um diese mit einem neuen Hashtag zu versehen und ihnen somit eine zusätzliche Bedeutung oder neue Interpretation zu verleihen (C. Oh et al., 2016). *Re-tagging* setzt natürlich ein vorheriges *Re-finding* voraus, da ältere Tweets erst lokalisiert werden müssen. Die Möglichkeit älteren Posts ein Update zu verleihen ist unter Social-Media-Plattformen eine einzigartige Funktion, tritt aber in ihren Analysen im Kontext des *#WHP*-Projektes bei etwa 35% aller Posts auf (C. Oh et al., 2016, S.1283). Neben den bereits erwähnten Möglichkeiten gibt die Praxis des *Re-tagging* den Nutzern die Möglichkeit, in Erinnerungen zu schwelgen, ältere Inhalte zu betrachten und das Bedürfnis nach Nostalgie zu befriedigen. Ähnliche Verhaltensweisen treten auf Instagram auch bei anderen Hashtags wie dem Hashtag *throwback thursday* (*#tbt*) auf. Auch auf Twitter finden Hashtags wie *#tbt* bzw. *#throwbackthursday* Verwendung (Leahey, 2014). Inwiefern die Einführung der neuen RT-Funktionalität bzw. die Kommentarfunktion auch zu einem Anstieg im Weiterleiten älterer Tweets geführt hat und ob vergleichbare Hashtag-Projekte auf Twitter Einfluss auf das *Re-finding* und RT-Verhalten haben, wird in der vorliegenden Arbeit nicht untersucht, sollte aber Gegenstand zukünftiger Studien sein.

### 3.3. Zusammenfassung: Twitter und Social-Media-PIM

Kapitel 3 hat gezeigt, dass Twitter die meist erforschte Social-Media-Applikation ist. Betrachtet man Twitter-Forschung aus der Perspektive der drei PIM-Hauptaktivitäten, so wird deutlich, dass die Perspektive des Auffindens von Tweets — oder wie Nutzer die Plattform verwenden, um an relevante Informationen zu gelangen — bereits sehr viel Aufmerksamkeit erfahren hat. Für die späteren Forschungsfragen sind speziell die Relevanzcharakteristika von Tweets von Interesse, die vor allem im Kontext des Ranking von Tweets besprochen werden, da diese auch vermeintlich Charakteristika von wiederfindenswerten Tweets darstellen könnten. Auch die Arbeiten von Kwak und Kollegen oder Wu und Kollegen, die Aussagen darüber treffen, wie stark sich Informationen auf der Plattform ausbreiten, welche Charakteristika hierzu beitragen und wie lange Tweets folglich von Relevanz für die Nutzenden sein können, sind für den vorliegenden Untersuchungskontext von Relevanz (Kwak et al., 2010; Wu et al., 2011). Wie Nutzer im Anschluss an das Auffinden der Informationen diese schließlich behandeln, ist so gut wie nicht untersucht. Die Verwaltung von Social-Media-Inhalten wird zunächst aus der Perspektive des *impression management* vorgestellt, wobei ein zentraler Punkt das Löschen (*regret deletion*) von eigenen gesendeten Inhalten darstellt. Löschverhalten setzt natürlich auch die Rückkehr bzw. das Wiederfinden dieser Tweets voraus, jedoch steht diese Perspektive nicht im Fokus der Arbeit. Eine interessante Perspektive eröffnen Studien, die die Wahrnehmung und Nutzung von Social-Media-Plattformen als persönliches Archiv vorstellen. Hauptmotivation hier ist die Präsentation des eigenen Ichs oder auch das Schwelgen in Erinnerung. Pinterest wird als PIM-Tool vorgestellt (Linder et al., 2014). Es wird interessant sein zu sehen, ob auch Twitter — trotz der Tatsache, dass es über einen sehr starken Echtzeitfokus verfügt — ähnliche Charakteristika eines Archivs aufzuweisen hat bzw. von Nutzern als Archiv wahrgenommen wird. Indizien, dass Nutzer Tweets aufbewahren und welche Strategien sie hierfür verwenden, finden sich in der Literatur nur an einzelnen Stellen. Lediglich die Studie von boyd, Golder und Lotan weist darauf hin, dass Nutzer das Bedürfnis haben, Tweets aufzubewahren, indem sie diese Weiterleiten (RT) und anschließend in den eigenen gesendeten Tweets lokalisieren (boyd et al., 2010). Gorrell und Bontcheva nehmen eine der ersten Studien vor, die auch das Favorisieren als Möglichkeit des Aufbewahrens vorstellt, indem es zum Setzen von Lesezeichen für Tweets benutzt wird (Gorrell & Bontcheva, 2016). Die erste Diskussion über die Frage, ob Nutzer Tweets wiederfinden, legen Teevan et al. in ihrer Twitter-Suchloganalyse vor. Sie tendieren jedoch dazu — anders als im Webkontext — die wiederholten Suchanfragen nicht als *Re-finding*-

Suchanfragen zu interpretieren, sondern aufgrund des Echtzeitfokus, als Monitoring-Verhalten zu bewerten (Teevan et al., 2011). Später relativieren sie ihre Aussage dahingehend, dass die Twitter-Suche das Bedürfnis des Wiederfindens nur sehr schlecht unterstützt und daher dieses Verhalten wenig beobachtet werden kann, was aber nicht bedeutet, dass es nicht existiert, sondern eher dazu führt, dass Nutzer andere Wiederfindensstrategien entwickeln (Dumais et al., 2014). In ihrer Tagebuchstudie zur Nutzung der Twitter-Suche stellen Elsweiler und Harvey fest, dass die meisten Suchanfragen ihrer Probanden durch Wiederfindensabsichten motiviert sind (Elsweiler & Harvey, 2015). Ob tatsächlich ein derart hoher Prozentsatz an Suchanfragen auf Wiederfinden abzielt, wird in der vorliegenden Arbeit zu klären sein.

# 4

## Forschungsdesign

Kapitel 4 stellt die Gestaltung des Forschungsprozesses näher vor. Bei der Organisation des Forschungsprozesses bzw. dem Forschungsdesign steht vor allem die Frage im Raum, wie die Validität — in diesem Sinne auch als Glaubwürdigkeit oder Authentizität zu interpretieren — der einzelnen Ergebnisse bzw. Erkenntnisse, sichergestellt werden kann. Aus der Perspektive der Sozialwissenschaften stellt Creswell acht Validierungsstrategien vor, wobei er die Strategie der Triangulation von Datenquellen — und Erkenntnissen aus diesen Quellen — als die am einfachsten zu implementierende und am häufigsten genutzte Variante zur Validierung von Forschungserkenntnissen charakterisiert (Creswell, 2014, S.201). Durch Validierungsstrategien wie Triangulation kann ein komplettes, objektives Bild des Forschungsgegenstandes geschaffen werden, was zu einem tieferen Verständnis des analysierten Gegenstandsbereiches führt. Dies könnte bei der Verwendung von nur einer Methode nicht erreicht werden (Fabritius, 1999, S.406). Als eine Ergänzung bzw. Weiterentwicklung der Triangulation kann das Forschungsprinzip der *Mixed Methods* interpretiert werden (Creswell, 2014). Das Verhältnis und die Charakteristika dieser beiden Validierungsstrategien bzw. Forschungsdesignprinzipien werden in Abschnitt 4.2 näher erläutert. Im Folgenden soll in einem kurzen Abriss die Bedeutung derartiger Forschungsdesignprinzipien für die Informationswissenschaft, im Speziellen für die Untersuchung des Informationsverhaltens, vorgestellt werden. Anschließend werden die beiden Forschungsdesignprinzipien Triangulation und *Mixed Methods* näher erläutert. Abschnitt 4.3 präsentiert schließlich das Forschungsdesign der Promotionsarbeit.

### 4.1. Triangulation und *Mixed Methods* in der Informationswissenschaft

Für die Untersuchung von Informationsbedürfnissen und dem Informationsverhalten haben Triangulationsstudien seit den neunziger Jahren verstärkt an Bedeutung gewonnen. Hannele Fabritius untermauert dies, indem sie eine Reihe von Arbeiten aus diesem Untersuchungskontext zitiert (Fabritius, 1999, S.406). Case und Given stellen fest, dass sich viele Studien zum Informationsverhalten lediglich auf eine singuläre Form der Datenerhebung bzw. -art konzentrieren und in Fällen, in denen mehrere Erhebungstechniken verwendet werden, lediglich offensichtlich zusammengehörige Verfahren wie Interviews und die Begutachtung von Artefakten wie Dokumente verwendet

werden (Case & Given, 2016, S.217). Weiter heben sie die Arbeit von Paul Solomon positiv hervor, welcher für die Untersuchung des *sense making* und Informationsverhalten bei der Erstellung von Arbeitsplänen in öffentlichen Einrichtungen eine Vielzahl an unterschiedlichen Erhebungs- und Auswertungsmethoden (Teilnehmende Beobachtung, Interviews, Dokumentanalyse) verwendet (Case & Given, 2016, S.266). Um die Verwendung bzw. Ergänzung von unterschiedlichen Methoden zu beschreiben, nutzt Case in der zweiten Ausgabe des Buchs *Looking for Information* noch keine Fachbegriffe wie Triangulation oder *Mixed Methods*, so dass dort keine Aussagen darüber getroffen werden können, wie stark derartige Forschungsdesignprinzipien bei der Untersuchung des Informationsverhaltens tatsächlich Verwendung finden (Case, 2007). In der aktuellen, vierten Ausgabe des Buches werden nun Arbeiten von Creswell und Fidel zum Kontext Mixed-Methods-Studien im Informationsverhalten angeführt (Case & Given, 2016, S.224ff.). Raya Fidel stellt fest, dass Mixed-Methods-Forschung in Studien zur Bibliotheks- und Informationswissenschaft nicht verbreitet ist und in der Literatur selten diskutiert wird (Fidel, 2008). Bei einer systematischen Inhaltsanalyse von 465 Artikeln verwenden nur 5% eine Herangehensweise in ihren Untersuchungen, die als *Mixed Methods* interpretiert werden können, jedoch wird der Begriff an sich gar nicht verwendet (Fidel, 2008). In einer weiteren Arbeit führen Peng, Nunes und Annansingh Gründe an, weshalb Mixed-Methods-Studien in der informationswissenschaftlichen Forschungslandschaft kaum zu finden sind und warum diese von der Community wenig bis gar nicht angenommen werden (Peng et al., 2011). Sie argumentieren, dass der hohe Planungsaufwand von Mixed-Methods-Studien — der ihrer Meinung nach nötig ist, um quantitative und qualitative Erkenntnisse zu integrieren — Forscher davon abhält, dieses Forschungsdesign zu implementieren (Peng et al., 2011). Erst in späteren Meta-Studien wird die Kategorie *Mixed Methods* auch als Studienart aufgeführt, wie in Abschnitt 6.2.1 zu lesen sein wird. Dort werden die Arbeiten von Heidi Julien et al. vorgestellt, welche in Meta-Studien zur Methodenverwendung in der Informationswissenschaft *Mixed Methods* als eine Methodenkategorie verwenden (Julien, 1996; Julien & Duggan, 2000; Julien et al., 2011; Julien & O'Brien, 2014).

## 4.2. Triangulation und Mixed-Methods-Forschungsdesigns im Detail

Das Konzept der Triangulation kann in allen Phasen des Forschungsprozesses eingesetzt werden. Denzin unterscheidet vier Varianten der Triangulation (Datentriangulation, Investigatortriangulation, Theorietriangulation, Methodentriangulation) (Schreier & Echterhoff, 2013, S.288). Meistens, wenn der Begriff der Triangulation fällt, ist die Methodentriangulation in der Variante Between-Methods-Triangulation gemeint, welche die „Verwendung unterschiedlicher Methoden zur Erfassung desselben Gegenstandes“ (Schreier & Echterhoff, 2013, S.288) meint. Margit Schreier definiert Triangulation wie folgt:

„Unter Triangulation im weiteren Sinne versteht man das Einnehmen unterschiedlicher Perspektiven auf denselben Forschungsgegenstand. In der Methodenliteratur und auch im vorliegenden Kontext bezeichnet Triangulation die Erhebung von Daten zu einem Gegenstand unter Anwendung von (mindestens) zwei verschiedenen Methoden“ (Schreier & Echterhoff, 2013, S.288).

Wie dem Zitat zu entnehmen ist, steht bei der Triangulation die Kombination mehrerer Methoden bei der Datenerhebung im Vordergrund. Bei der Triangulation können sowohl zwei qualitative als auch zwei quantitative Methoden miteinander kombiniert werden. Anders ist dies beim Mixed-Methods-Design der Fall. Bei *Mixed Methods* kann die Integration von Methoden, Verfahren und Techniken während der Forschung als Abgrenzungskriterium zur Triangulation interpretiert werden. Im Falle von *Mixed Methods* müssen verpflichtend sowohl qualitative als auch quantitative



Methoden miteinander kombiniert werden (Kuckartz, 2014, S.30). Jedoch beschränkt sich dies nicht auf den Prozess der Datenerhebung. Auch Auswertungsmethoden können das Methodenspektrum erweitern und somit zu einem Mixed-Methods-Ansatz führen (Schreier & Echterhoff, 2013, S.291). Studien im Mixed-Methods-Design sind häufig auch Triangulationsstudien, jedoch nicht jede Studie, die es sich zum Ziel setzt, Triangulation zu realisieren, ist auch eine Mixed-Methods-Studie. Kuckartz weist darauf hin, dass eine Vielzahl an unterschiedlichen Definitionen von *Mixed Methods* existieren, wobei unterschiedliche Schwerpunktsetzungen in den Definitionen zu erkennen sind (Kuckartz, 2014). Er beruft sich dabei auf einen Artikel von Johnson, Onwuegbuzie und Turner, welche bei dem Vorhaben einen eigenen Definitionsansatz zu formulieren, 19 verschiedene Definitionen von *Mixed Methods* diskutieren (Johnson et al., 2007). Kuckartz selbst liefert eine Arbeitsdefinition, die eine „methodische Akzentuierung“ vornimmt und *Mixed Methods* als Kombination und Integration von qualitativen und quantitativen Methoden versteht (Kuckartz, 2014, S.33). Besonderer Fokus ist auf den letzteren Teil der Definition zu legen, in der es heißt: „Die Integration beider Methodenstränge, d.h. von Daten, Ergebnissen und Schlussfolgerungen, erfolgt je nach Design in der Schlußphase des Forschungsprojektes oder bereits in früheren Projektphasen“ (Kuckartz, 2014, S.33). Die genannte Form der Ergebnisdarstellung entspricht auch der Darstellung, wie sie in der vorliegenden Arbeit verfolgt wird. Zunächst werden die Erkenntnisse der einzelnen Studien separat dargestellt, jedoch wird in Abschnitt 7.1 eine Integration aus dem Erkenntnisstand der Literatur und den Ergebnissen der beiden Studien vor dem Hintergrund der gestellten Forschungsfragen erfolgen.

Neben unterschiedlichen Definitionsansätzen gibt es im Bereich *Mixed Methods* auch zahlreiche Differenzierungen, die sowohl auf Kombinationen der Methodenart (qualitativ vs. quantitativ) zurückgehen, als auch auf den zeitlichen bzw. inhaltlichen Bezug der einzelnen Studien. So existieren auch Designvarianten mit Bezeichnungen wie Multimethod-Design, Multimethod-Studie oder Monomethod-Studie. Im Sinne von Schreier können all diese Bezeichnungen bzw. Arten als Varianten von *Mixed Methods* verstanden werden, wobei *Mixed Methods* den Oberbegriff darstellt (Schreier & Echterhoff, 2013, S.292). Ergänzend zu diesen Begrifflichkeiten bestehen auch unterschiedliche Systematisierungsansätze von Mixed-Methods-Design. Morse unterscheidet neun verschiedene Mixed-Methods-Designvarianten in einer Matrix mit den beiden Dimensionen *Kriterium der Gewichtung*, also ob qualitative und quantitative Daten gleich gewichtet werden oder einer der beiden Daten ein höheres Gewicht zukommt und *Kriterium der Reihenfolge*, also ob beide Datenarten gleichzeitig erhoben werden oder eine sukzessive Datenerhebung stattfindet (Schreier & Echterhoff, 2013, S.299). Creswell et al. ergänzen diese beiden Dimensionen um die Dimensionen der *Integration*, also zu welchem Zeitpunkt die Integration der Daten bzw. Ergebnisse erfolgt und der Dimension der *theoretischen Perspektive*, also welchen theoretischen Rahmen das gesamte Forschungsdesign besitzt (Creswell et al., 2003).

Creswell unterscheidet zwischen drei Grundarten des Mixed-Methods-Design (Creswell, 2014, S.220):

- *Convergent Parallel Mixed Methods*: Die Erhebung und Auswertung der quantitativen bzw. qualitativen Daten erfolgt parallel. Die Erkenntnisse werden zusammengeführt, in Beziehung gesetzt, verglichen und schließlich interpretiert.
- *Explanatory Sequential Mixed Methods*: Die Datenerhebung erfolgt dort sequentiell. Zunächst werden quantitative Daten gesammelt. Anschließend erfolgt eine Nachfolge-Studie qualitativer Natur. Diese Designform wird auch als Vertiefungsdesign bezeichnet.
- *Exploratory Sequential Mixed Methods*: Aufbauend auf einer qualitativen Studie wird eine quantitative Studie durchgeführt. Eine Methode dient folglich zur Entwicklung der nächsten (Kuckartz, 2014, S.58). Als bestes Beispiel kann hier die Erstellung und Durchführung einer Fragebogenstudie genannt werden, deren Items auf einer zuvor durchgeführten Interview-

studie basieren. Diese Designform wird auch als Verallgemeinerungsdesign bezeichnet.

Neben diesen drei Grundarten stellt Creswell folgende drei fortgeschrittene Varianten vor: *Embedded Mixed Methods*, *Transformative Mixed Methods* und *Multiphase Mixed Methods* (Creswell, 2014, S.227-233). Eine nähere Auseinandersetzung mit diesen Varianten würde aber zu weit führen. Teddlie und Tashakkori erstellen die sog. „Methods-Strands Matrix: A Typology of Research Designs Featuring Mixed Methods“ (Teddlie & Tashakkori, 2006). In ihr werden mehrere Studiendesignvarianten verortet, wobei in Zelle vier die Designansätze angeführt werden, die auch bei Creswell genannt werden. Das Forschungsdesign der vorliegenden Arbeit entspricht der von Teddlie und Tashakkori beschriebenen Variante eines Sequential-Mixed-Designs (Teddlie & Tashakkori, 2006, S.22). Diese Form des MM-Design ist, folgt man den Autoren, gut dafür geeignet, um explorativen Studien einen Forschungsrahmen zu bieten (Teddlie & Tashakkori, 2006, S.22). Charakteristisch für das sequentielle Design ist die zeitlich versetzte, fortlaufende Abfolge in sich geschlossener Einzelstudien und somit die Tatsache, dass die Daten unabhängig voneinander gesammelt und ausgewertet werden. Erkenntnisse einer früheren Studie können jedoch das Design der Folgestudie beeinflussen. Zudem erfolgt in einem finalen Schritt — der Phase der Meta-Inferenz — eine Referenz der einzelnen Ergebnisse aufeinander. Im Sinne einer Triangulation bzw. der Komplementarität werden Übereinstimmungen zwischen Resultaten der Einzelstudien gefunden und durch Elaboration und Illustration der Ergebnisse der einen Methode, kann ein besseres Verständnis der Ergebnisse der anderen Methode erreicht werden et vice versa (Kuckartz, 2014, S.58).

Jede der durchgeführten Studien besteht aus vier grundlegenden Stufen oder Phasen. Nach Teddlie und Tashakkori sind die einzelnen Phasen wie folgt definiert (Teddlie & Tashakkori, 2006, S.16):

- Konzeptualisierungsphase (*Conceptualization stage*): Phase der Konzepterarbeitung, in der Forschungszweck und Forschungsfragen formuliert werden
- Experimentelle Phase I (*Methodological stage*): Phase der Methodenauswahl in der konkrete Beobachtungen vorgenommen und Daten gesammelt werden
- Experimentelle Phase II (*Analytical stage*): Phase der Datenanalyse
- Schlussfolgerungsphase (*Inferential stage*): Phase der Ergebnispräsentation in der aufbauend auf den Ergebnissen abstrakte Erklärungsansätze oder Theorien formuliert werden

Insofern beide Studien sowohl qualitative als auch quantitative Methoden der Datenerhebung und Datenauswertung vereinen, stellt das Gesamtdesign einen Spezialfall dar, da auch die Einzelstudien einen gewissen Mixed-Methods-Charakter besitzen.

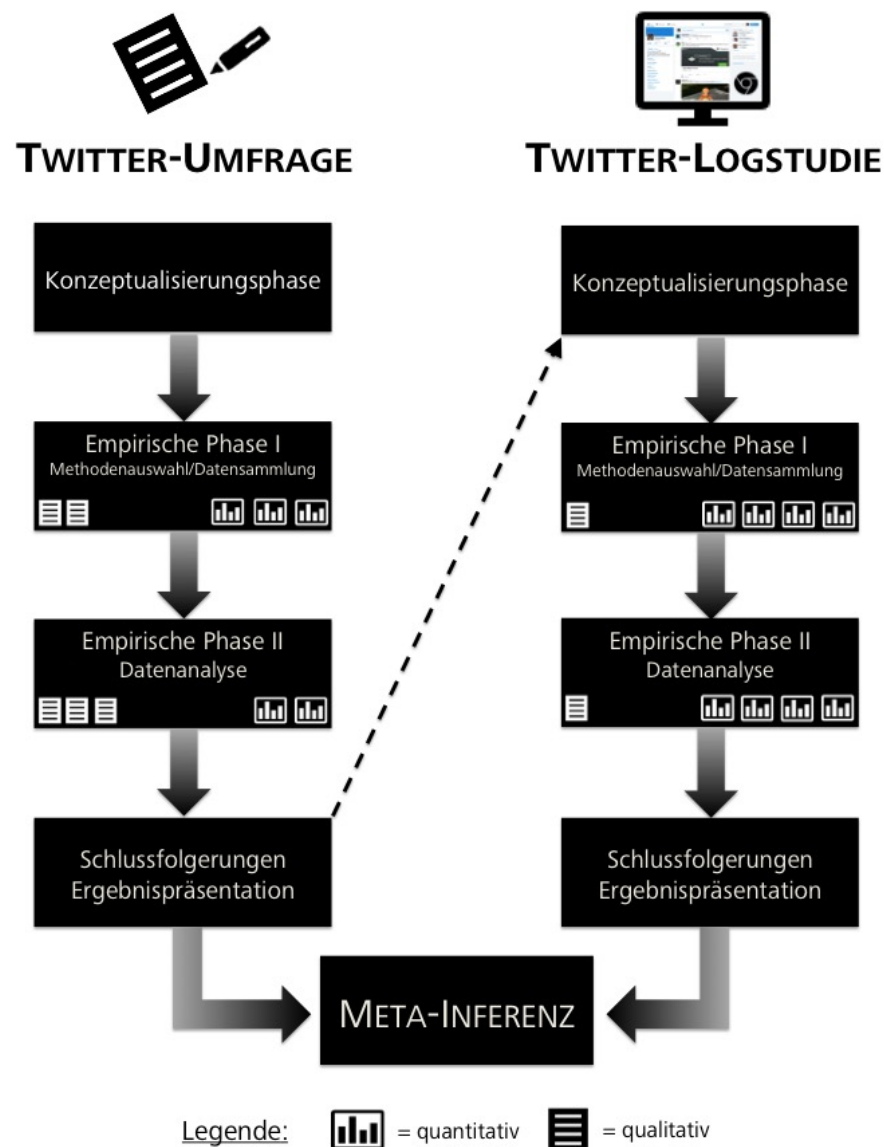
## 4.3. Das Forschungsdesign im Überblick

Abbildung 4.1 zeigt das Forschungsdesign im Überblick. Wie bereits erwähnt, sieht das Sequential-Mixed-Design vor, dass die Datenerhebung sowie die Datenanalyse für beide Studien voneinander unabhängig durchgeführt werden und erst zur Phase der Meta-Inferenz eine Zusammenführung der Erkenntnisse aus beiden Studien erfolgt. Im Folgenden werden die wesentlichen Phasen des Forschungsdesigns kurz vorgestellt und eine Zuordnung zwischen den Phasen der Studien und Kapiteln der Arbeit vorgenommen.

### 4.3.1. Twitter-Umfrage

#### Konzeptualisierungsphase

Erste Ideen Personal-Information-Management-Praktiken im Kontext von Social-Media-Applikationen



**Abbildung 4.1.:** Visualisierung des Forschungsdesigns auf der Basis des Sequential-Mixed-Design nach Teddlie und Tashakkori (Teddlie & Tashakkori, 2006)

zu untersuchen, stoßen die Konzeptualisierungsphase der Twitter-Umfrage an. Zunächst wurde eine breit angelegte Literaturlauswertung hinsichtlich des Forschungsvorhabens durchgeführt, welche in Kapitel 2 zum Thema PIM entlang der PIM-Praktiken Finden, Aufbewahren und Wiederfinden durchgeführt wurde und in Kapitel 3 zum Thema Social-Media-Nutzung mit Fokus auf PIM und Twitter-Verhalten getätigt wurde. Aufbauend auf dieser Auswertung werden erste Ideen zu Forschungsfragen formuliert, welche in Abschnitt 5.1 aufgeführt sind.

#### **Empirische Phase I: Methodik**

Die Auswertung der Literatur führte zur Erkenntnis, dass das Forschungsvorhaben noch nicht auf einem Niveau ergründet ist, dass es erlaubt hätte, als initiale Studie bereits eine hypothesentestende Studie oder Evaluation durchzuführen. Ziel der initialen Studie ist es, einen ersten Überblick über PIM-Verhaltensweisen auf Twitter zu gewinnen und zu verstehen, wie Technologie, die als Echtzeit-Informationsnetzwerk konzipiert ist, genutzt wird, wenn PIM-Probleme, wie Wiederfinden, auftreten. Der durch die explorative Fragebogenstudie gewonnene *high-level* Überblick dient dazu, konkretere Fragestellungen zu formulieren. In den Abschnitten 5.2.1 und 5.2.2 werden die Konzeption sowie die einzelnen Bestandteile des Fragebogens erläutert. Die Gewinnung der Studienteilnehmer für die Fragebogen-Umfrage — also Rekrutierungsstrategien und Samplingverfahren — werden in Abschnitt 5.3 erläutert.

Mithilfe der Icons in Abbildung 4.1 wird der Fokus der Datensammlung visualisiert. Für die empirische Phase I ist hier zu erkennen, dass der Schwerpunkt bei der Datenerhebung auf der quantitativen Seite liegt, da der Fragebogen weitaus weniger offene Freitextfragen als auf Likert-Skalen basierende geschlossene Fragen enthält. Dennoch werden in der Twitter-Umfrage sowohl quantitative als auch qualitative Daten erhoben.

### **Empirische Phase II: Auswertung**

Die empirische Phase II ist geprägt von der Datenauswertung. Für die Auswertung der qualitativen Daten wird in Abschnitt 5.4 eine ausführliche Diskussion der zwei bekanntesten Auswertungsverfahren *Grounded Theory* und qualitative Inhaltsanalyse durchgeführt. Zudem erfolgt eine Beschreibung des von Margrit Schreier entworfenen Werkzeugkastenmodells der Inhaltsanalyse (vgl. Abbildung 5.9), welches zur konkreten Datenanalyse herangezogen wird bzw. dessen Phasen bei der Datenauswertung adaptiert werden (vgl. hierzu Abschnitt 5.4.3). Zur Analyse der quantitativen Daten werden zum einen die Signifikanzprüfung mithilfe des nicht-parametrischen *Bootstraps* (vgl. Abschnitt 5.7.1) sowie zum anderen die logistische Regression (vgl. Abschnitt 5.7.8.1) als Auswertungsverfahren vorgestellt und schließlich in den Abschnitten 5.7, bei der Darstellung der quantitativen Ergebnisse zu Aufbewahrens- und Wiederfindensstrategien bzw. in Abschnitt 5.7.8, bei der Analyse der Einflussfaktoren auf das Frustrationsniveau der Nutzer beim Wiederfinden von Tweets, angewendet. Die Anzahl der Icons in Abbildung 4.1 verdeutlicht, dass bei der Datenanalyse der Fokus auf der qualitativen Datenanalyse lag. Hierfür gibt es mehrere Gründe. Zum einen, da eine qualitative Datenanalyse — in diesem Fall wurde der Prozess zweimal bei der Analyse der Favorisierungsgründe sowie der Analyse der Wiederfindensgründe durchlaufen — zeitintensiv und aufwändig ist. Zum anderen, weil qualitative Daten sehr wertvoll sind und eine Menge hochwertiger Einblicke in das Verhalten der Nutzer bieten können. Für das Ziel der Twitter-Umfrage — Erlangung eines Überblicks über PIM- bzw. Wiederfindensverhalten auf hoher Ebene und insbesondere dessen Motivationsgründe — bietet diese Art der Analyse eine optimale Grundlage für den weiteren Forschungsverlauf bzw. um in der Phase der Meta-Inferenz qualitative Ergebnisse durch quantitative Daten zu untermauern.

### **Schlussfolgerungen und Ergebnispräsentation**

Die Twitter-Umfrage betrachtet Wiederfindensverhalten unter folgenden vier Gesichtspunkten:

1. Das Kodierschema der Favorisierungsgründe aus Kapitel 5.6.3, welches 25 unterschiedliche Motive aufdeckt, das Feature zu nutzen und zudem beschreibt, welche Rolle diese Funktion für das PIM-Verhalten bzw. das Wiederfinden von Tweets besitzt.
2. Die Abschnitte 5.7.2 bis 5.7.7 beleuchten Aufbewahren und Wiederfinden aus quantitativer Perspektive und zeigen, wie Einstellungen seitens der Nutzer das Wiederfindensverhalten beeinflussen.
3. Aufbauend auf den quantitativen Ergebnissen zu Wiederfindenshäufigkeit und -strategien sowie Faktoren, die diese beeinflussen können, liefert das logistische Regressionsmodell aus Kapitel 5.7.8 Informationen darüber, wann Wiederfinden als frustrierend eingeschätzt wird

und welche Faktoren dazu beitragen können, Frustration zu vermeiden oder zu reduzieren.

4. Das Kodierschema der Wiederfindensgründe aus Abschnitt 5.8 fördert zahlreiche Gründe für das Wiederfinden von Tweets zu Tage.

Diese Ergebnisse werden in den jeweiligen Abschnitten separat vorgestellt und beschrieben. Im Zuge der Meta-Inferenz werden die einzelnen Ergebnisse erneut zusammengebracht, um im Sinne der Triangulation die verschiedenen Facetten der Forschungsfragen bzw. Untersuchungsgegenstände, welche durch die unterschiedlichen Methoden eruiert werden, ergründen zu können. Somit wird ein vollständiges, kohärentes Bild von Wiederfinden erschaffen (Schreier & Echterhoff, 2013, S.289). Bereits die Einzelergebnisse der Twitter-Umfrage beeinflussen die Konzeptualisierungsphase der Twitter-Logstudie, was durch den Pfeil in Abbildung 4.1 angedeutet ist.

### 4.3.2. Twitter-Logstudie

#### Konzeptualisierungsphase

Die Twitter-Umfrage ist aufgrund des qualitativen Auswertungsfokus darauf ausgelegt, einen Überblick über PIM/Wiederfindensverhalten zu erlangen und vor allem Motivation und Beweggründe seitens der Nutzer zu sammeln. Aussagen von Probanden können durchaus von einer gewissen Subjektivität bzw. einem kognitiven *Bias* geprägt sein, der immer dann auftritt, wenn Nutzer ihre Absichten oder eigenes Verhalten beschreiben müssen.<sup>1</sup> Um die Stärke dieses *Bias* bzw. den Grad der Subjektivität der Ergebnisse abschätzen zu können und auch tatsächliches Verhalten der Nutzer zu beobachten, wurde als Anschlussstudie eine lebensnahe, objektive Logdatenerhebung und -auswertung durchgeführt, die Nutzer bei ihrer tatsächlichen Interaktion mit Twitter über eine gewisse Zeit aufmerksam betrachtet. Die Erkenntnisse der Twitter-Umfrage erlauben es, die in Abschnitt 5.1 formulierten Leitfragen zur Twitter-Umfrage zu konkretisieren und somit auf den lebensnahen, quantitativen Kontext der Twitter-Logstudie anzupassen. Die Leitfragen der Twitter-Logstudie sind in Abschnitt 6.1 formuliert.

#### Empirische Phase I: Methodik

Die Methode der Logdatenerhebung bzw. -auswertung wird in Abschnitt 6.2 vorgestellt. Kapitel 6.3.1 erklärt die wesentlichen Bestandteile des für die Logstudie konzipierten und eigens implementierten Erhebungsinstruments, der Chrome Webbrowser-Erweiterung (vgl. Abschnitt 6.3.1). Der Abbildung 4.1 ist zu entnehmen, dass die Twitter-Logstudie sowohl bei der Datenerhebung als auch bei der Datenauswertung einen starken quantitativen Fokus besitzt. Dennoch werden auch bei der Twitter-Logstudie qualitative Daten erhoben, was mithilfe der Experience-Sampling-Methode geschieht, die in Abschnitt 6.3.2 — sowohl methodisch als auch in ihrer technischen Umsetzung — näher charakterisiert wird.

#### Empirische Phase II: Auswertung

Aufgrund des starken quantitativen Schwerpunkts werden für die Datenanalyse verschiedenste statistische Analyseverfahren verwendet. Die Charakterisierung des allgemeinen Twitter-Verhaltens (vgl. Abschnitt 6.4.3) wird vor allem durch Methoden der deskriptiven Statistik in Form von Visualisierungen, Tabellen und der Angabe von Lagemaßen durchgeführt. Das tatsächliche Wiederfindensverhalten wird ebenfalls mithilfe deskriptiver Statistik analysiert. Zusätzlich werden, um Gruppenunterschiede festzustellen, verschiedenste Tests wie z. B. der Chi-Quadrat-Test verwendet. Beim Vergleich des Wiederfindensverhaltens auf Sessionebene (vgl. Kapitel 6.4.4.4) wird ebenfalls das in Kapitel 5.7.1 eingeführte Verfahren des Signifikanztests mithilfe des nicht-parametrischen Bootstraps angewendet. Ein besonderes Verfahren der Datenanalyse wird in Kapitel 6.4.5 ange-

<sup>1</sup>Für eine Liste an möglichen ergebnisverfälschenden Biases siehe z. B. die Arbeit von Choi und Pak (Choi & Pak, 2005)

wendet. Hier wird das maschinelle Lernverfahren *Random Forest* dazu genutzt, um in einer Reihe von Experimenten unter der Verwendung von verschiedenen Feature-Gruppen Wiederfindensziele vorherzusagen. Eine Erläuterung dieses Verfahrens erfolgt in Abschnitt 6.4.5.1.

#### **Schlussfolgerungen und Ergebnispräsentation**

Die Twitter-Logstudie bietet zahlreiche Erkenntnisse über Twitter-Verhalten im Allgemeinen und Re-finding-Verhalten im Speziellen:

- Erstmalige Charakterisierung des allgemeinen Twitter-Verhaltens. Es werden erste Erkenntnisse über typisches Twitter-Verhalten präsentiert z. B. die Länge von Twitter-Sitzungen oder die Bedeutung von Unterseiten neben der *Timeline* (vgl. Kapitel 6.4.3).
- Zeitliche Dimension des Wiederfindens und der Lebenszyklus bestimmter Tweets für Individuen (vgl. Abschnitt 6.4.4.3).
- Re-finding-Verhalten wird auf vielfältige Art und Weise beschrieben sowohl auf abstrakter Sessionebene als auch mit Einzelereignissen (vgl. Abschnitte 6.4.4.4 bis 6.4.4.8).
- *Feature engineering* führt zu einer Reihe an Eigenschaften, die Wiederfinden charakterisieren, da Wiederfinden bzw. Wiederfindensziele mit einer angemessenen Genauigkeit vorhergesagt werden können. Dies wird in Abschnitt 6.4.5 näher vorgestellt.

### **4.3.3. Meta-Inferenz**

Kapitel 7 stellt, *per definitionem* des Forschungsdesigns, das Kapitel dar, in welchem die Meta-Inferenz vorgenommen wird. Hierbei werden die Erkenntnisse der beiden Studien mit den bestehenden Kenntnissen aus der Literatur zusammengebracht und diskutiert. Hierbei erfolgt eine *Triangulation*, um die verschiedenen Facetten des Gegenstandsbereichs Re-finding zu erfassen, bzw. — im Idealfall — ein sich ein ergänzendes und somit vollständiges Bild von Wiederfindensverhalten auf Twitter zu beschreiben (Schreier & Echterhoff, 2013, S.289).

# 5

## Studie I: Twitter-Umfrage

Kapitel 5 präsentiert die erste der beiden Studien. Sie analysiert PIM-Praktiken und Wiederfindensverhalten auf Twitter mithilfe einer webbasierten Umfrage. Die Stichprobe der Probanden wurde als *convenience sample* via verschiedener Rekrutierungsstrategien, u.a. durch die Verwendung eines Crowdsourcing-Panels, zusammengestellt. Aufgrund des hohen qualitativen Anteils der Studie nimmt die vergleichende Diskussion der qualitativen Methoden *Grounded Theory* (GT) und *qualitative Inhaltsanalyse* einen großen Raum im Kapitel ein. Die Methodendiskussion führt schließlich zur Beschreibung der für die Auswertungen verwendeten Methodik, die einen Methoden-Mix darstellt und auf dem Werkzeugkastenmodell nach Margrit Schreier basiert (vgl. Abschnitt 5.4.2). Insgesamt umfasst das Kapitel drei Schwerpunkte:

1. Die quantitative und qualitative Charakterisierung der Nutzung der Favoriten-Funktion und ihre Rolle für PIM-Praktiken insbesondere Re-finding.
2. Die quantitative Beschreibung des Aufbewahrens- und Wiederfindensverhaltens und die Analyse von Faktoren, die ein frustrierendes Erlebnis beim Wiederfinden von Tweets steigern oder vermindern können.
3. Die Erstellung und Beschreibung einer Systematik, die die Bandbreite an unterschiedlichen Wiederfindensgründen erfasst und diskutiert.

Eine Interpretation und Einordnung der in diesem Kapitel entstehenden Erkenntnisse erfolgt schließlich unter Berücksichtigung der bereits bestehenden Forschung zum Wiederfindensverhalten in anderen Domänen und den Erkenntnissen der quantitativen Twitter-Logstudie aus Kapitel 6 in Kapitel 7, dem Kapitel der Meta-Inferenz des Sequential-Mixed-Forschungsdesigns.

Im Sinne einer publikationsgestützten Dissertation sind einige Inhalte dieses Kapitels bereits einem Bewertungsverfahren unterzogen und in Form von Publikationen in den Tagungsbänden der jeweiligen Konferenzen veröffentlicht.

Die Ideen aus Abschnitt 5.6 — in dem die Rolle der Favoriten-Funktion für das PIM-Verhalten auf Twitter erläutert wird — wurden auf der *The 8th International AAAI Conference on Weblogs and Social Media* (ICWSM'14) vorgestellt und im dazugehörigen Tagungsband veröffentlicht (Meier et al., 2014).

Abschnitt 5.7.8, welcher die Einflussfaktoren auf das Frustrationsniveau bei Wiederfindensaktionen auf Twitter in Form des Logistischen Regressionsmodells beschreibt, wurde bereits im Tagungs-

band zur IliX'14 (*Proceedings of the 5th Information Interaction in Context Symposium*) veröffentlicht (Meier & Elsweiler, 2014b).

Auch die Forschungs- bzw. Leitfragen, welche die gesamte Promotionsarbeit — aber vor allem die Twitter-Umfrage — motiviert haben, wurden im Rahmen des *Doctoral Consortium* der *5th Information Interaction in Context Symposium* als Vortrag präsentiert und im dazugehörigen Tagungsband veröffentlicht (Meier & Elsweiler, 2014a).

## 5.1. Leitfragen der Twitter-Umfrage

Ziel der Umfrage ist es, einen ersten Eindruck davon zu erlangen, wie verbreitet Aktivitäten mit PIM-Bezug (Aufbewahren — Organisieren — Wiederfinden) auf Twitter sind. Kapitel 3.2 zeigt, dass zum Forschungsbeginn nur die Informationssuche im Fokus von Studien zu PIM-Verhalten auf Twitter stand. Nur wenige Arbeiten thematisieren die Frage nach dem Langzeitinformationsgehalt von Tweets für individuelle Nutzer und die Rückkehr zu diesen Informationen (boyd et al., 2010; Teevan et al., 2011; Elsweiler & Harvey, 2015; Gorrell & Bontcheva, 2016). Im Folgenden werden die wichtigsten Leitfragen der Twitter-Umfrage aufgeführt. Viele dieser Fragen erfahren später in den Leitfragen zur Twitter-Logstudie noch eine Konkretisierung (vgl. Abschnitt 6.1).

### **Twitter als Archiv: In welchem Ausmaß ereignen sich Aktivitäten mit PIM-Bezug?**

Andere Social-Media-Applikationen wie Facebook oder Pinterest werden von einem gewissen Teil ihrer Nutzer als Archiv wahrgenommen (vgl. Abschnitt 3.2.2.1). Inwiefern auch bei Twitter eine solche Wahrnehmung besteht, ist eine zentrale Fragestellung. Wie hoch ist die Frequenz von PIM-Aktivitäten? Wie häufig haben Nutzer das Bedürfnis, Tweets aufzubewahren bzw. zu Tweets zurückzukehren? Wie häufig wollen Nutzer Tweets bewusst wiederfinden? Wie häufig handelt sich dabei nicht um Tweets aus der *Timeline*, sondern um eigene gesendete Tweets? Gibt es Situationen, in denen nicht nur einzelne Tweets von Interesse sind, sondern das Bedürfnis des Nutzers nur gestillt werden kann — wie im Falle von E-Mail Re-finding (Elsweiler et al., 2008) — wenn mehrere Tweets wiedergefunden werden? Obwohl hier nicht als separate Leitfrage aufgeführt, wird in der Twitter-Umfrage eine klare Differenzierung zwischen Aufbewahrensverhalten (*keeping*) und Re-finding vorgenommen.

### **Welche Aufgaben und Bedürfnisse motivieren Nutzer dazu, Tweets wiederzufinden?**

Die Gründe und Motive, die Nutzer zur Rückkehr zu Tweets bzw. zum Wiederfinden motivieren, bilden einen wesentlichen Teil der Arbeit und helfen bei der Beantwortung der Frage nach den Charakteristika des Wiederfindens von Tweets. Einige Dimensionen wurden bereits im vorherigen Abschnitt genannt. So kann es nicht nur das Ziel des Nutzers sein, Tweets aus der *Timeline* wiederzufinden, sondern auch eigene gesendete Tweets könnten vom Nutzer wiedergefunden werden wollen. Welche Rolle spielen Zeit, Tweetinhalt oder Sender eines Tweets als Motivation für das Wiederfinden? All dies sind mögliche Einflussfaktoren auf Wiederfindensmotive.

### **Welche Aufbewahrens- und Wiederfindensstrategien werden verwendet?**

Wie gehen Nutzer vor, wenn sie Tweets wiederfinden wollen? Verwenden sie interne Funktionen, oder gehen sie sogar so weit und suchen nach Tweets in externen Speichern oder mithilfe von Suchmaschinen? Gibt es unterschiedliche Wiederfindensstrategien für verschiedene Motive? Welchen Einfluss besitzt die Wiederfindenshäufigkeit auf die verwendeten Strategien und die Verwendung von Twitter-Funktionalitäten? Welche Rolle spielt die Einstellung, die ein Proband in Bezug auf Twitter hat auf dessen Verhalten? All dies sind mögliche Fragestellungen, die in diesem Kontext von Relevanz sind und im Fragebogen erhoben werden. Obwohl die formulierten Fragen stets auf Wiederfinden ausgerichtet sind, sind sie identisch auch auf das Aufbewahrensverhalten übertragbar.



**Welche UI-Features oder -Optionen bietet Twitter für PIM-Aktivitäten?**

Diese Frage stellt eine Konkretisierung der Frage nach den allgemein verwendeten Aufbewahrens- und Wiederfindensstrategien dar, indem sie zusätzlich die Auswahlmöglichkeiten, die auf Twitter-UI-Features basieren, in den Vordergrund stellt. Einige Twitter-Funktionen bieten laut Literatur oder Hinweisen aus Blogs und Webseiten die Möglichkeit, sie für PIM-Zwecke einzusetzen (vgl. Abschnitt 3.2.2.2). Um zu prüfen, ob diese Features auch tatsächlich für die angedachten Zwecke — z. B. zum Aufbewahren von Tweets — genutzt werden, wurden die Probanden der Twitter-Umfrage nach der Nutzung der Funktionen für den jeweiligen Kontext befragt. Bei der Frage nach Aufbewahrensstrategien liegt der Fokus auf der Nutzung der Favoriten-Funktion, da sie die primäre Funktionalität der Plattform darstellt, die ein Abspeichern und somit das Verwalten von Tweets erlaubt und einen späteren, erneuten Zugang (Wiederfinden) erleichtert und beschleunigt. Darüber hinaus stellt sich grundsätzlich die Frage, wie Nutzer über PIM und Twitter denken, also wie sie die Funktionalität der Plattform für diese Aktivitäten einschätzen und ob sie Wiederfinden generell als schwierig zu lösendes Informationsbedürfnis erachten.

**Wie können Social-Media-Applikationen gestaltet werden, um das Nutzerbedürfnis *Re-finding* zu unterstützen?**

Die Analyse des Nutzungsverhaltens ist nicht nur zum Selbstzweck gedacht. Es sollen auch Verbesserungsmöglichkeiten zur Unterstützung des Wiederfindensverhalten herausgearbeitet werden. Letztendlich ist dies nicht alleine durch die Überblicks-Umfrage lösbar. Erst wenn alle Facetten der Verhaltensweisen erkannt sind, können auch adäquate Verbesserungsvorschläge gemacht werden. In Kapitel 8 wird es schließlich mögliche Designvorschläge gegeben, die aufzeigen, dass Wiederfindensverhalten unterstützt werden kann und wie dies möglich ist. Da aber von Beginn an auch diese Perspektive auf das Thema bedacht wurde, wird diese Forschungsfrage der Vollständigkeit halber hier angeführt.

Die Hauptleitfragen für die Twitter-Umfrage spiegeln sich schließlich auch in der Konzeption bzw. den Bestandteilen des Fragebogens wider, der insgesamt aus den drei großen Bereichen Favorisierungsverhalten, Aufbewahrensverhalten und Wiederfindensverhalten zusammengesetzt ist. Die Konzeption des Fragebogens wird im Folgenden beschrieben.

## 5.2. Fragebogendesign und Fragebogengliederung

In den Human-, Sozial- und Kommunikationswissenschaften ist die schriftliche Befragung ein gängiges und weitverbreitetes Verfahren. Es ist deshalb nur konsequent, dass zahlreiche Lehr- oder Methodenbücher ein Kapitel bzw. mehrere Abschnitte zur adäquaten Durchführung dieser Methodik enthalten (Bortz & Doering, 2006; Fowler & Cosenza, 2008; Kirchhoff et al., 2010; Kallus, 2010; Porst, 2014). Sämtliche Regeln des guten Fragebogendesigns hier aufzuführen und zu erläutern, würde den Rahmen dieses Abschnitts sprengen. Die Konzeption eines Fragebogens kann in die Phase der Makroplanung und die Phase der Mikroplanung unterschieden werden (Bortz & Doering, 2006). Die Überlegungen bzw. die Literatur, welche bei diesen Planungsschritten von Relevanz sind, werden in den jeweiligen Abschnitten 5.2.1 sowie 5.2.2 beschrieben. Die Erstellung des Fragebogens war von einem iterativen Prozess geprägt. So wurden mehrmals Prototypen des Fragebogens erstellt, die durch informelle Pre-Tests überprüft wurden, um missverständliche Fragen zu identifizieren und zu verbessern. Der Fragebogen wurde eigens für diese Umfrage in einer webbasierten Variante als Single-page-Webanwendung in JavaScript/JQuery implementiert. Für das Frontend-Design wurde das Framework *Bootstrap* verwendet. Mit Hilfe eines PHP-Backend-Skripts wurden die Fragebogendaten an eine MySQL-Datenbank gesendet. Die webbasierte Variante wurde auf einem Server der Universität Regensburg gehostet und erhielt eine Subdomain

der Informationswissenschafts-Website<sup>1</sup>, um als seriöser Fragebogen wahrgenommen zu werden. Das Vorgehen bei der Crowdsourcing-Variante wird in Abschnitt 5.3.2 beschrieben. Im Folgenden werden Makro- sowie Mikroplanung des Fragebogens näher erläutert.

### 5.2.1. Makroplanung des Fragebogens

Der Fragebogen besteht aus 37 Items, die in fünf Fragenblöcke gegliedert sind. Um dem Prinzip einer sinnvollen Fragensukzession zu entsprechen, wurden die Fragenblöcke so gegliedert, dass sie einem allgemeinem PIM-Verhalten entsprechen (Aufbewahren vor Wiederfinden) und die Häufigkeit des Verhaltens stets vor konkreten Verhaltensweisen abgefragt wurde (Porst, 2014, S.137). Die thematische Ausrichtung sowie die Verteilung der Items war wie folgt:

1. Demographie und Twitter-Nutzung (10 Items)
2. Favorisierens-Verhalten (1 bzw. 10 Item(s))
3. Aufbewahrens-Verhalten (1 bzw. 5 Item(s))
4. Wiederfindens-Verhalten (1 bzw. 10 Item(s))
5. Zusammenfassung (2 Items)

Bis auf die Fragenblöcke zur Demographie und Zusammenfassung sind alle Fragenblöcke mit Filterfragen versehen, die je nach Antwort der Probanden dazu führte, dass diese weitere Items zum jeweiligen Themenblock erhalten, oder zum nächsten Themenblock weitergeleitet werden.

### 5.2.2. Mikroplanung des Fragebogens und Frage-Antwort-Typen

Um den Fragebogen so verständlich wie möglich zu gestalten, wurde die Anzahl der Frage-Antwort-Kombinationsarten auf drei Varianten beschränkt. Die Befragten fanden somit häufig Fragen mit einem identischen Antwortmodus vor. Ein Umstand, der in der Fragebogenliteratur empfohlen wird (Kallus, 2010, S.18). Die Fragebogenitems zur Messung der Frequenz bzw. des Ausmaßes von Verhaltensweisen wurden stets als geschlossene Fragen formuliert, da „[f]ür die Erkundung konkreter Sachverhalte die [...] Frageform besser geeignet“ ist (Bortz & Doering, 2006, S.254). Als Antwortmöglichkeit für die Frequenzfragen wird eine siebenstufige Likert-Skala gewählt. Kallus argumentiert, dass sich sieben Stufen bei Likert-Skalen etabliert haben (Kallus, 2010, S.40). Fowler und Cosenza führen Studien an, die Evidenz dafür liefern, dass Antwortskalen mit bis zu sieben Punkten die Qualität der Messung verbessern (Fowler & Cosenza, 2008, S.151). Jeder Skalenpunkt wurde mit einem Label versehen. Man nennt dies eine *verbalisierte* Skala (Porst, 2014, S.79). Verfügt jeder Skalenpunkt über ein eigene Beschriftung, führt dies zu weniger kognitiver Belastung bei der befragten Person, jedoch muss der Fragebogenentwickler für entsprechend aussagekräftige Labels sorgen (Porst, 2014, S.81). Zudem führen verbalisierte Skalen zu konsistenteren Antworten bei den Probanden (Fowler & Cosenza, 2008, S.151). Die Label-Beschriftungen wurden aus einer Umfrage von Morris, Teevan und Panovich zur Nutzung von Twitter als soziales Frage-Antwort-System übernommen, welche sich dort als passend erwiesen haben (Ringel Morris et al., 2010). Abbildung 5.1 zeigt exemplarisch die Frage nach der Häufigkeit des Sendens von Tweets aus dem Frageblock zur Demographie und allgemeinen Twitter-Nutzung und der entsprechende Sieben-Punkt-Antwortskala mit den verbalisierten Skalenpunkten *never* bis *multiple times per day*. Insgesamt beinhaltet der Fragebogen elf Fragen, die die Probanden nach der Häufigkeit von bestimmten Verhaltensweisen befragten und diese Antwortskala besitzen.

<sup>1</sup><http://twittersurvey.iw.ur.de>

1.7 How often do you update your status message (post a tweet)?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Abbildung 5.1.:** Verbalisierte Sieben-Punkt-Antwortskala wie sie bei Fragen zur Häufigkeit von Verhaltensweisen genutzt wird

Die Fragenblöcke, welche die Studienteilnehmer nach Verhaltensweisen in konkreten Situationen befragten, wurden vom Aufbau her stets identisch formuliert. Lediglich die konkreten Fragen bzw. Antwortmöglichkeiten werden entsprechend des jeweiligen PIM-Kontextes variiert. Als übergeordnete Frage stand stets die Formulierung: *Which of the following best applies?*, wobei diese von mehreren konkreten Fragen (Itemkern) begleitet wird. Dies soll exemplarisch anhand von Abbildung 5.2 verdeutlicht werden. Wie der Abbildung zu entnehmen ist, ist jede Frage mit einer fünf-stufigen, verbalisierten Likert-Skala in Form einer Zustimmungsskala (*strongly disagree* bis *strongly agree*) versehen, wodurch die ProbandInnen ihre Ablehnung bzw. Zustimmung zu den als Fragen formulierten Verhaltensweisen tätigen. Eine ungerade, fünf-stufige Skala birgt immer die Gefahr, dass die mittlere Kategorie von den Befragten als Fluchtkategorie genutzt wird. Eine ungerade Skala kann jedoch auch positive Seiten besitzen. Sie ist stets eine Option für den Fall, dass der Nutzer die Frage nicht beantworten, oder sich einfach für keine Seite der Skala entscheiden kann (Porst, 2014, S.83). Es wurde eine ungerade Skala gewählt, da negative Effekte, wie sie bei geraden Skalen auftreten — wie *item no response* oder die zufällige Wahl irgendeines Punktes — vermieden werden sollten (Porst, 2014, S.84).

Als dritte Fragevariante wurden offene Fragen formuliert, die die Probanden zu Gründen/Motiven für bestimmte Verhaltensweisen befragten. Offene Fragen sollten vor allem dann zur Informationsgewinnung für ein bestimmtes Thema genutzt werden, wenn über den Befragungsgegenstand noch wenig bekannt ist, bzw. die Dimensionen von Antwortmöglichkeiten zum Zeitpunkt der Untersuchung nicht überschaubar sind (Porst, 2014, S.67). Beim Fragenformat der offenen Frage wurde nicht nur eine einzelne Frage gestellt, sondern der ProbandIn ein Duo aus Frage und Arbeitsauftrag präsentiert. Zunächst war die Frage allgemein formuliert, worauf auch eine generelle Antwort erwartet wird. Dann wurden Nutzer dazu aufgefordert, einen konkreten Beispiel-Tweet zu suchen, der die Befragten zum Nachdenken anregt und eine konkrete Verhaltenssituation evozieren soll. Offene Fragen fordern häufig ein gutes Erinnerungsvermögen seitens der ProbandInnen. Bortz charakterisiert Fragen, die ein gutes Erinnerungsvermögen bei den Probanden voraussetzen, wie die Rekonstruktion von Tagesabläufen, als problematisch (Bortz & Doering, 2006, S.255). Aus diesem Grund wurde versucht, ein konkretes Beispiel — in diesem Fall durch einen konkreten Tweet, den Nutzer z. B. aus ihrer Favoritenliste auswählen — als Gedächtnisstütze zu nutzen. Auch bei Morris, Teevan und Panovich wurde versucht, den *recall bias* zu überwinden, indem konkrete Tweets von den Probanden gefordert waren (Ringel Morris et al., 2010). Im Folgenden werden die einzelnen Fragenblöcke des Fragebogens näher vorgestellt. Der Fragebogen ist in Anhang A.1 zu finden.

#### **Begrüßungsseite und Teilnehmerhinweise**

Vor Beginn des Fragebogens wurden die Teilnehmer über Ursprung und Zweck des Fragebogens aufgeklärt. Zudem wurden einige weitere Hinweise zum Ablauf der Erhebung gegeben. Abbildung 5.3 zeigt den Begrüßungs-Screen des Fragebogens.

#### **Fragenblock: Demographie und allgemeine Twitter-Nutzung**

Neben klassischen demographischen Fragen zu Geschlecht, Alter und Schulausbildung (Fragen 1.1—1.3) wurden speziell Hintergrundinformationen zur Twitter-Nutzung (Fragen 1.4—1.7) abgefragt. So wurden die ProbandInnen nach ihrer gewöhnlichen Zugangssituation (mobile Appli-

### 3.3 Which of the following best applies?

When you want to preserve a Tweet for later, do you...

	strongly disagree	disagree	neutral	agree	strongly agree
tend to mark it as favourite?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to retweet it, so you can look it up in your own posted Tweets when you need to?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
use some additional store (e.g. browser bookmark, software (Pocket, Evernote...), E-mail inbox, Word document, paper printout, RSS Feed)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to do nothing special with such Tweets?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Abbildung 5.2.:** Verbalisierte Fünf-Punkt-Zustimmungsskala wie sie bei Fragen nach konkreten Verhaltensweisen genutzt wird

kation, Twitter-Website etc.) befragt und die Häufigkeit mit der sie Tweets versenden (vgl. Abbildung 5.1) ermittelt. Die Frage nach den allgemeinen Account-Informationen (Fragen 1.8—1.10) — hierfür wurden die ProbandInnen gebeten sich bei Twitter einzuloggen — schließt die Sammlung demographischer Informationen ab. Der Abschnitt wurde bewusst kurz gehalten, da keine Informationen darüber vorliegen, ob bzw. wie derartige demographischen Aspekte das PIM-Verhalten auf Social-Media-Plattformen beeinflussen und auch Analysen auf der Basis von demographischen Charakteristika nicht im Fokus der Arbeit liegen.

#### Fragenblock: Favorisierensverhalten

Dem Fragenblock zum Favorisierensverhalten wird eine Filterfrage (Frage 2.1) vorgeschaltet, die Nutzer, welche die Favoriten-Funktion nicht kennen, davor bewahrt, weitere Fragen zu diesem Themenblock zu erhalten. Nutzer, die die Favoriten-Funktion kennen, erhalten einen weiteren Frageblock, welcher als erstes die Frage nach der Anzahl der bereits favorisierten Tweets stellt (Frage 2.2). Hierfür wurde via Twitter-API ein kleines, webbasiertes Programm implementiert, welches nach der Eingabe des Twitter-Nutzernamens (z. B. @meier\_flo) die Anzahl der favorisierten Tweets anzeigt. Anschließend werden Fragen zum Favorisierensverhalten gestellt wie die Häufigkeit der Funktionsnutzung, Gründe für die Nutzung sowie Häufigkeit und Gründe für das Entfavorisieren (Fragen 2.3—2.6). Dieselben Fragen wurden sowohl in Bezug auf Tweets von anderen Personen als auch in Bezug auf eigene gesendete Tweets gestellt (Fragen 2.7—2.8).

#### Fragenblock: Aufbewahrensverhalten

Dem Frageblock zum Aufbewahrensverhalten ist ebenfalls eine Filterfrage vorgeschaltet. Im Anschluss an diese Frage wurden Nutzer zunächst nach Gründen bzw. Motiven für das Aufbewahren von Tweets befragt. Im Anschluss daran werden explizit Methoden zur Aufbewahrung aufgelistet, zu denen Nutzer ihre Ablehnung oder Zustimmung äußerten (Frage 3.3). Hierfür wurden auf der Basis der Literatur vier Antwortmöglichkeiten formuliert:

- Das Favorisieren, da es in der Funktionsbeschreibung von Twitter als Bookmarking-Funktion für Tweets beschrieben ist. Auch in der breiten Öffentlichkeit — betrachtet man Blog-Einträge zum Thema — wird Favorisieren als Lesezeichen-Funktion wahrgenommen (Stone, 2006).
- Das Weiterleiten (RT) von Tweets, welches in der Studie zum RT-Verhalten von boyd, Golder und Lotan, als eine Funktion zum Aufbewahren bzw. erleichterten Wiederfinden von Tweets charakterisiert wurde (boyd et al., 2010).
- Die Nutzung eines — in diesem Fall nicht näher definierten — externen Speicherortes. Hierfür wurde eine Vielzahl an Möglichkeiten aufgezählt. In Studien zum PIM-Verhalten bei Web-

**Twitter Survey**

**Introduction**

Dear User,

the purpose of this survey is to analyze Information Behaviour on Twitter. The questionnaire is voluntary and the data collected is strictly confidential. All participants will NOT be identified and you have the option not to answer a particular question. All data and measurements obtained from this research study will be stored confidentially. Only researcher will have access to view any data collected during this research.

This survey should only take about 10 minutes of your time. Questions to specific aspects of your Twitter account will occur. So for your convenience open a separate browser window and log in to [Twitter](#).

During the survey please don't use the back button of your browser. JavaScript has to be enabled.

☐ I confirm, that I read the introduction and I am over 18 years of age.

**Continue with Survey**

**Abbildung 5.3.:** Begrüßungs-Screen der webbasierten Variante der Umfrage

seiten oder E-Mails wird diese Form des *keeping behaviour* immer wieder beobachtet, weshalb sie auch bei Social-Media-Nachrichten genutzt werden könnte (Bruce et al., 2004a,b; Whitaker et al., 2011).<sup>2</sup>

- Als letzte Alternative wird keine Methode im eigentlichen Sinne aufgeführt, sondern die Option gegeben, keine explizite Methode zu nennen, aber dennoch der Meinung zu sein, einen aufbewahrenswerten Tweet gelesen zu haben.

Auch hier wurden dieselben Fragen in Bezug auf eigene gesendete Tweets gestellt (Fragen 3.4-3.5).

#### **Fragenblock: Wiederfindensverhalten**

Analog zu den anderen Frageblöcken steht vor dem Frageblock zum Wiederfindensverhalten eine Filterfrage (Frage 4.1). Anschließend wurden die Probanden nach einem Tweet befragt, den sie wiederfinden wollten und nach Motiven für das Wiederfinden dieses Tweets (Frage 4.2-4.3). Frage 4.4 befragte Probanden nach detaillierten Methoden/Verhaltensweisen beim Wiederfinden. Hierfür wurden folgende Alternativen geboten:

- Das Lokalisieren des Tweets in der eigenen *Timeline* via Scrolling. Die Frage hierfür lautete: *tend to search in your own Timeline*. Dies ist leider eine ambige Formulierung, da die eigene *Timeline* auch als Sammlung der eigenen gesendeten Tweets interpretiert werden kann, was dem Pendant zum Aufbewahren via Weiterleiten (RT) entsprechen würde.
- Als Pendant zum Aufbewahren von Tweets via Favorisieren: das Lokalisieren des Tweets in der Favoritenliste.
- Nutzung der Twitter-Suche und Absetzen einer Suchanfrage (Query). Diese Auswahl wurde aufgrund der thematischen Nähe zur Studie von Teevan und Kollegen aufgenommen, welche das Verhalten bei der Suchanfrageformulierung von Twitter-Nutzern untersuchen und Re-finding via Suche als seltenes, aber mögliches Verhalten identifizieren (Teevan et al., 2011).
- Nutzung einer Websuchmaschine wie Google, Bing oder Yahoo zur Lokalisierung des Tweets.
- Als Pendant zur Aufbewahrung von Tweets in externen Speichern: Das Lokalisieren des Tweets in einem externen Speicherort.

Diese Fragen wurden ebenfalls in Bezug auf eigene gesendete Tweets gestellt (Fragen 4.5-4.7). Neben diesen Verhaltensfragen wurde den ProbandInnen ein Frageblock zu konkreten Wiederfin-

<sup>2</sup>In der Tat gab einer der Nutzer in einem der offenen Fragen an, *Packratius* für Aufbewahrungszwecke zu verwenden.

densgründen vorgelegt (Frage 4.8). Abschließend erhielten die Befragten zwei Fragen zur zeitlichen Dimension des Wiederfindens. Frage 4.9 stellte die Frage nach einer gewöhnlichen Wiederfindensperiode in Form einer verbalisierten fünfstufigen Likert-Skala, wohingegen Frage 4.10 die Nutzer nach einem Maximum für das Wiederfinden in Form einer offenen Freitextfrage befragte.

#### **Fragenblock: Zusammenfassung**

Zum Abschluss des Fragebogens wurde den Probanden die Aufgabe gestellt, die Möglichkeiten, die Twitter zum Aufbewahren und Wiederfinden von Tweets bietet, auf einer fünfstufigen Skala zu bewerten (Frage 5.1). Zudem wurden Nutzer befragt, wie häufig sie bei Wiederfindensaktionen auf Twitter frustriert worden sind.

### **5.3. Samplingverfahren und webbasierte Rekrutierungsstrategien für Umfragen**

Als Zielgruppe für die Umfrage kann die Gesamtheit aller Twitter-Nutzenden gelten, wobei es unrealistisch ist, eine Vollerhebung durchzuführen und Daten von der gesamten Population zu erhalten, so dass eine geeignete bzw. *repräsentative* Stichprobe ausgewählt werden muss. In der Statistik unterscheidet man probabilistische (zufällige) und nicht-probabilistische Verfahren bei der Stichprobenauswahl, wobei erstere immer den Vorzug erhalten sollten, da sie eine größere Aussagekraft besitzen (Bortz & Doering, 2006, S.402). In der Praxis ist es jedoch oft aus finanziellen, infrastrukturellen und organisatorischen Gründen nicht möglich, Zufallsstichproben zu erstellen, weshalb auf Ad-hoc-Stichproben oder Stichproben durch Bequemlichkeitsauswahl (*convenience sample*) zurückgegriffen wird. Stichproben können dann als repräsentativ bezeichnet werden, wenn sie in ihren Merkmalen bzw. Merkmalskombinationen mit der Population vergleichbar sind (Bortz & Doering, 2006, S.396). Es stellt sich die Frage, ob dies für den vorliegenden Fall überhaupt möglich bzw. notwendig ist. Folgende Punkte sind hierbei anzuführen:

- Da es sich im vorliegenden Fall um eine explorative Untersuchung handelt, die vielmehr darauf abzielt, Hypothesen zu generieren, als bestehende Hypothesen zu überprüfen, ist die Auswahl der Probanden sowie die Größe der Stichprobe weitestgehend unerheblich (Bortz & Doering, 2006, S.70).
- Ein wesentlicher Bestandteil des Fragebogens orientiert sich an qualitativer Forschung (Gründe für das Favorisieren (vgl. Abschnitt 5.6), Gründe für das Wiederfinden von Tweets (vgl. Abschnitt 5.8)), indem verbales Material interpretiert wird. Bei dieser Art der Erhebung ist eine Standardisierung des Untersuchungsvorgangs nur in geringem Umfang nötig (Bortz & Doering, 2006, S.297).
- Inferenzstatistische Tests können auch dann durchgeführt werden, wenn keine Zufallsstichprobe im strengen Sinne erzeugt wurde. Es handelt sich nicht um eine mathematische Voraussetzung, sondern lediglich um eine Vorgabe, die für die Interpretation/Generalisierung der Ergebnisse eine Rolle spielt. Für jede Bequemlichkeitsstichprobe ist *ex post* eine Inferenzpopulation rekonstruierbar (Bortz & Doering, 2006, S.401).

Die Auswahl von Versuchsteilnehmern für Umfragen oder Experimente basierend auf der Einfachheit ihrer Verfügbarkeit wird als *convenience sampling* oder Bequemlichkeitsstichprobe bezeichnet. Mit der Ankunft des partizipativen *Web 2.0* und Social-Software-Applikationen haben auch die Möglichkeiten der Zusammenstellung von Stichproben via webbasierter Rekrutierungsstrategien zugenommen. Grundsätzlich ist der Frage, ob und wie gut sich Onlinequellen zur Sammlung von Umfragedaten eignen, die generelle Frage vorgelagert, ob online- bzw. webbasierte Verfahren zur Sammlung von Umfragedaten identisch oder vergleichbar sind mit den Resultaten eines physischen

Fragebogens, den die Versuchspersonen mit einem Stift vor Ort auszufüllen haben. Für Behrend und Kollegen besteht in der Forschung Konsens darüber — sie belegen dies, indem sie eine Reihe an Studien zitieren — dass diese beiden Verfahren größtenteils äquivalente Charakteristika in Bezug auf psychometrische Gütekriterien, Antworttendenzen (*social desirability*) oder Vollständigkeit der Daten aufweisen (Behrend et al., 2011). Andere Faktoren wie Antwortrate oder Absprungrate können sehr stark von der demographischen Zusammensetzung des Marktplatzes oder Panels abhängen, wie Lozar-Manfreda und Kollegen anhand des Amazon-Mechanical-Turk-Marktplatzes zeigen (Lozar-Manfreda et al., 2008). Die Bandbreite an Rekrutierungsstrategien ist so gut wie unbegrenzt, da jedes Mittel zur Kommunikation mit anderen Personen dazu genutzt werden kann, um diese zur Teilnahme an einer Umfrage zu rekrutieren. Ob auf Foren-ähnlichen Plattformen wie *Reddit* oder Social-Media-Applikationen wie *Facebook*, *Twitter* oder *LinkedIn*, Links zu Umfragen können auf vielfältige Weise positioniert werden. So existieren auch Studien, die hinterfragen, wie gut Social-Media-Applikationen dazu genutzt werden können, um *snowball sampling* zu betreiben (Baltar & Brunet, 2012).

Neben diesen häufig kostenlosen Möglichkeiten, Umfragen zu verbreiten und ein *convenience sample* zu gewinnen, besteht auch die Möglichkeit webbasierte Umfrage-Software wie SurveyMonkey<sup>3</sup>, Qualtrics<sup>4</sup> oder — wie auch im vorliegenden Fall — TellWut<sup>5</sup> zu nutzen. Diese Plattformen bieten häufig nicht nur die benötigte Software zur Erstellung und Verbreitung von webbasierten Umfragen, sondern verfügen auch über ein eigenes Umfrage-Panel das — jedoch oft gegen Gebühren — auch gleich zur Rekrutierung von Probanden genutzt werden kann. Neben hohen Preisen besitzt dieses Verfahren weitere Grenzen, wie der Umstand, dass diese Panels nicht über die nötige Zielgruppe verfügen, die angesprochen werden soll. Bei stark abgegrenzten Stichproben, also Stichproben, die ein bestimmtes demografisches Profil der Teilnehmer verlangen, sind nicht die nötigen Probanden registriert. Auch im vorliegenden Fall handelt es sich bei der Rekrutierung der Studienteilnehmer um eine scharf abgegrenzte Stichprobe, da nur Twitter-Nutzer als Versuchspersonen in Frage kommen. Die Nutzung von Panels der Umfrage-Software-Anbieter ist vergleichbar mit der Nutzbarmachung von sog. Crowdsourcing-Plattformen oder -Marktplätzen zur Rekrutierung von Studienteilnehmern, ein Vorgehen welches im Folgenden näher erläutert wird.

### 5.3.1. Die Verwendung von Crowdsourcing zur Rekrutierung von Probanden für Fragebogenstudien

Der Term *Crowdsourcing* ist ein vom Autor und Journalisten Jeff Howe geprägter Neologismus aus den Begriffen *crowd* und *outsourcing*. In einer ursprünglichen *white paper version* einer Begriffsdefinition bezeichnet Howe Crowdsourcing als „[...] the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call.“ (Howe, 2006). Je nachdem aus welcher Perspektive bzw. wissenschaftlicher Fachrichtung Crowdsourcing thematisiert wird, stehen unterschiedliche Charakteristika dieses Phänomens im Vordergrund. Meier identifiziert drei Perspektiven des Crowdsourcings (Meier, 2012): (i) Die ökonomische Perspektive, in welcher Crowdsourcing als wirtschaftliches Instrument oder Geschäftsmodell betrachtet wird. (ii) Die informatische Perspektive, welche die technologischen Voraussetzungen des Crowdsourcings beschreibt, die die Möglichkeit des partizipatorischen Mitwirkens an einer Sache via Web-Interfaces erst ermöglichen. (iii) Die soziokulturelle Perspektive, welche die Konsequenzen für die Gesellschaft, den Arbeitsmarkt und auch den *digital divide* beschreibt (Meier, 2012). Zudem werden in der Literatur sehr vie-

<sup>3</sup><https://www.surveymonkey.de/>

<sup>4</sup><https://www.qualtrics.com/de/>

<sup>5</sup><https://www.tellwut.com/>





gen. Micro-Task-Plattformen wie *Amazon Mechanical Turk* (MTurk) haben sich in der Wissenschaft für die Zwecke der Teilnehmerrekrutierung etabliert. Crowdsourcing-Plattformen werden in den unterschiedlichen wissenschaftlichen Disziplinen für Onlinestudien eingesetzt, von den Sozialwissenschaften über Psychologie und Medizin bis hin zu Informatik und Software Engineering, um große Mengen an Daten zu validieren, evaluieren oder annotieren (M. Swan, 2012; Palmer & Stickland, 2016; Mao et al., 2017). Dabei werden diese Plattformen nicht nur genutzt, um Versuchspersonen für Studien zu gewinnen, sondern sie sind auch selbst Gegenstand der Forschung. Im Folgenden werden Studien vorgestellt, die die Tauglichkeit von Crowdsourcing-Plattformen für die Rekrutierung von Probanden für Umfragen analysieren und diskutieren. Es wird ersichtlich werden, dass folgenden Aufgabenfaktoren eine Rolle spielen:

- Die Art und Formulierung der Aufgabe
- Die Höhe der Bezahlung für ihre Erfüllung
- Dauer, Einfachheit und Interessantheit der Aufgabe

In einem sehr frühen Artikel zur Datenqualität von Umfragedaten basierend auf MTurk-Stichproben äußern sich Buhrmester, Kwang und Gosling positiv über diese Art der Rekrutierungsform (Buhrmester et al., 2011). Sie vergleichen ein MTurk-Sample mit einem herkömmlichen *internet sample* und stellen dabei fest (Buhrmester et al., 2011, S.3):

- MTurk-Probanden sind demographisch diverser charakterisiert.
- Die Teilnehmerrate hängt stark von der Bezahlung sowie Art und Dauer der Aufgabe ab. Bei kurzen Umfragen (< 5 Min.) und hoher Bezahlung (50 cents) können sehr schnell sehr viele Probanden rekrutiert werden.
- Ist die Bezahlung angemessen, hat ihre Höhe keinen Effekt auf die Datenqualität.
- Die Datenqualität entspricht den psychometrischen Standards. Die Reliabilität der Daten ist ähnlich hoch, wie sie via traditioneller Umfragen erreicht werden kann.

Im Fazit ihrer Studie prognostizieren sie, dass — für den Fall, dass die Qualität der Daten konstant bleibt — sich Crowdsourcing als Hauptrekrutierungsstrategie für Umfragen in der Psychologie und den Sozialwissenschaften etablieren kann.

Um das Potential von Crowdsourcing-Plattformen für die Rekrutierung von Probanden für webbasierte Umfragen zu eruieren, vergleichen Behrend, Sharek, Meade und Wiebe eine Stichprobe zusammengesetzt aus PsychologiestudentInnen einer amerikanischen Universität mit einer Crowdsourcing-Stichprobe, die über MTurk rekrutiert wird (Behrend et al., 2011). Dabei stellen sie fest, dass die Crowdsourcing-Stichprobe ähnliches Verhalten offenbart wie die Stichprobe der StudentInnen. Das Crowdsourcing-Sample besitzt leichte Nachteile in Bezug auf die Datenqualität — so waren die Werte für soziale Erwünschtheit (*social desirability*) dort höher — aber aufgrund der demographischen Vielfalt — es konnten mehr ältere, erfahrenere und Personen mit diversen sozioökonomischen Hintergründen rekrutiert werden — besitzt es eine höhere Reliabilität. Behrend und Kollegen schließen ihre Studie mit dem Kommentar: „Our findings show that the use of crowdsourcing can be a potentially viable resource for researchers wishing to collect survey data on many types of organizational phenomena“ (Behrend et al., 2011, S.810).

Smith, Roster, Golden und Albaum vergleichen ein herkömmliches Online-Panel mit mehreren MTurk-Stichproben (in den USA beheimatete MTurker vs. nicht in den USA beheimatete MTurker) (Smith et al., 2016). Sie stellen fest, dass die Ergebnisse der drei Stichproben signifikante Qualitätsunterschiede (Teilnehmerintegrität und Datenqualität) aufweisen und ziehen den Schluß, dass die Wahl eines Online-Panel-Anbieters kritisch ist, da sie wesentliche Auswirkungen auf Stichprobenzusammenstellung, Datenqualität und Umfrageergebnis haben kann (Smith et al., 2016).

Antoun, Zhang, Conrad und Schober vergleichen vier Onlinequellen (*Craigslist*<sup>6</sup>, *Google AdWords*<sup>7</sup>, *Facebook* und *MTurk*) zur Rekrutierung von Studienteilnehmern für eine webbasierte Umfrage. Die zu gewinnende Stichprobe kann dabei als stark abgegrenzt bezeichnet werden, da nur iPhone-Nutzer angesprochen werden sollten (Antoun et al., 2016). Sie charakterisieren die beiden Quellen *Craigslist* und *MTurk* als sog. Pull-in-Strategien, da dort potentielle Umfrageteilnehmer aktiv nach Aufgaben/Mini-Jobs — wie eben der Teilnahme an einer Umfrage — suchen. *Google Adwords* und *Facebook* charakterisieren sie hingegen als Push-out-Strategien, da diese Rekrutierungsverfahren Werbung für die Teilnahme an der Studie an Nutzer senden, die sich eigentlich in einer normalen Nutzungssituation befinden und nicht aktiv danach suchen (Antoun et al., 2016, S.2). Sie vergleichen diese vier Quellen bzw. die beiden Rekrutierungsstrategien hinsichtlich folgender Kriterien: Kosteneffizienz, demographische Heterogenität und Datenqualität. Antoun und Kollegen stellen fest, dass Pull-in-Rekrutierungsquellen kosteneffizienter sind (*Craigslist*: umsonst, *MTurk*: \$1.67 pro Teilnehmer vs. *Google AdWords*: \$11,48, *Facebook*: \$30,22), eine höhere Datenqualität besitzen (längere Antworten bei Freitextfragen, weniger leere Antworten, mehr persönliche Details wie Einkommensdaten) und mehr Teilnehmer pro Tag zu rekrutieren vermögen (Antoun et al., 2016, S.8). Push-out-Strategien hingegen rekrutierten eine in Bezug auf demographische Charakteristika heterogenere Stichprobe, da die Pull-in-Strategien hauptsächlich jüngere, weibliche Teilnehmerinnen am unteren Rand der Einkommensgrenze mobilisieren (Antoun et al., 2016, S.12). Schließlich äußern sich Antoun et al. über die Leistungsfähigkeit von Online-Panels wie *Survey Monkey* oder *Qualtrics*, wobei sie deren Charakteristika ähnlich derer von Pull-in-Strategien sehen. Sie schließen mit der Empfehlung, stets unterschiedliche Varianten an Rekrutierungsstrategien- bzw. -quellen zu testen, da es keine One-size-fits-all-Rekrutierungsmöglichkeit gibt.

Auch Casler, Bickel und Hackett vergleichen bei einem verhaltenspsychologischen Fragebogen mehrere online-basierte Rekrutierungstechniken. Es handelt sich um ein Crowdsourcing-Sample via *MTurk*, sowie ein Sample, welches via Nachrichten auf verschiedenen Social-Media-Applikationen (Twitter, Reddit, Facebook) rekrutiert wird. Diesen beiden online-basierten Varianten stellen Casler et al. ein gewöhnliches Sample aus Universitätsstudenten gegenüber, welche den Fragebogen zudem nicht als webbasierte Variante sondern als Papier-Bleistift-Variante erhalten (Casler et al., 2013). Die demographischen Charakteristika der drei Samples sind sehr unterschiedlich, wobei die *MTurk*-Stichprobe auf signifikante Weise sozioökonomisch und ethnisch diverser ist. Die Qualität der Daten ist hingegen beinahe identisch, so dass sie eine Rekrutierung von Studienteilnehmern via *MTurk* als valide — in Bezug auf demographische Charakteristika sogar überlegene — Form der Rekrutierung charakterisieren (Casler et al., 2013). In einer Arbeit von Rouse weist dieser darauf hin, dass bei einem Persönlichkeitsfragebogen die *MTurk*-Stichprobe signifikant weniger verlässliche Daten geliefert hat als eine herkömmliche Stichprobe (Rouse, 2015). Zudem stellt er fest, dass die Länge des Fragebogens sowie die Höhe der Bezahlung nicht die Ursachen für die Datenqualität sind (Rouse, 2015).

Es wird offensichtlich, dass die Frage nach der Qualität der erhaltenen Daten eine entscheidende Frage bei der Nutzung von Crowdsourcing-Plattformen wie *MTurk* ist. Für Umfrageteilnehmer kann der schnelle monetäre Gewinn im Vordergrund stehen (Ipeirotis et al., 2010; Gadiraju et al., 2015). Auch hierzu existieren separate Studien. Peer, Vosgerau und Acquisti vergleichen zwei qualitätssichernde Maßnahmen: *MTurk*-Arbeiter mit hohem Ansehen (95% korrekt bearbeitete Aufgaben) und die Nutzung von sog. *attention check questions* (ACQs), welche das Engagement der Arbeiter vorab testen sollen, also eine Art Goldstandard darstellen und somit garantieren, dass nur engagierte Arbeiter den Fragebogen ausfüllen (Peer et al., 2014). Sie stellen fest, dass die Reduzierung auf *MTurk*-Arbeiter mit hoher Reputation genügt, um die Qualität der Daten hoch zu halten. Gadiraju, Kawase, Dietze und Demartini analysieren die Verhaltensmuster von *MTurk*-

<sup>6</sup><http://www.craigslist.org/about/sites#US>

<sup>7</sup><https://www.google.com/adwords/>

Arbeitern bei der Teilnahme an einer webbasierten Umfrage, um zwischen vertrauenswürdigen und unseriösen Arbeitern differenzieren zu können (Gadiraju et al., 2015). Sie zeigen, dass Goldstandards nicht die nötige Sicherheit aufbringen, um eine hohe Datenqualität zu gewährleisten, da bestimmte Gruppen von unseriösen Arbeitern diese Sicherheitsmaßnahme umgehen (Gadiraju et al., 2015). Aufbauend auf ihren Erkenntnisse formulieren Gadiraju und Kollegen Richtlinien für die Erstellung von Umfragen für MTurk (Gadiraju et al., 2015).

Fasst man die letzten Abschnitte zusammen, welche allesamt Studien vorstellen, die die Eignung von Crowdsourcing — insbesondere der Mikro-Task Plattform *Amazon Mechanical Turk* — zur Rekrutierung von Probanden für Umfragen thematisieren, so lässt sich durchaus ein Konsens dahingehend feststellen, dass Crowdsourcing als Rekrutierungsstrategie geeignet ist. Im Detail sprechen folgende Punkte für die Wahl dieser Rekrutierungsstrategie.

- Crowdsourcing ist eine bequeme, flexible und schnelle Rekrutierungsmethodik.
- Im Vergleich zu anderen Online-Panels, wie *Survey Monkey*, ist die Rekrutierung via Crowdsourcing eine preiswerte Alternative.
- In Bezug auf demographische Charakteristika können mithilfe von Crowdsourcing weitaus diversere, heterogene Stichproben rekrutiert werden.
- In den meisten Studien konnten die psychometrischen Gütekriterien der Stichproben die Erwartungen der Forschergruppen erfüllen.

Jedoch existieren auch Einschränkungen dieser Rekrutierungstechnik, die bei ihrer Nutzung beachtet werden müssen:

- Die Beschaffenheit der Umfrage (Länge, Komplexität, Thema) sowie die Bezahlung der Probanden kann die Qualität der Daten nachhaltig beeinflussen (auch wenn dies nicht immer festgestellt werden konnte (vgl. z. B. Rouse (Rouse, 2015))). Qualitätssichernde Maßnahmen oder Richtlinien für die Erstellung von web-basierten Umfragen für Crowdsourcing-Plattformen existieren und können für die Erstellungen der eigenen Umfragen genutzt werden (Gadiraju et al., 2015).
- Auch Crowdsourcing-Plattformen stellen keine Universallösung für die Rekrutierung von Studienteilnehmern dar. Wie von Antoun und Kollegen beschrieben, gibt es keine One-size-fits-all-Lösung, so dass immer mehrere Varianten getestet werden sollten, respektive es ratsam ist, klassische Samples zum Vergleich heranzuziehen (Antoun et al., 2016).

Im Folgenden wird die für das Crowdsourcing-Sample benutzte Umfrage-Plattform Tellwut, welche eine auf Umfrage spezialisierte Crowdsourcing-Plattform darstellt, näher charakterisiert.

### 5.3.2. Die Crowdsourcing-Umfrage-Plattform Tellwut

Tellwut<sup>8</sup> ist eine *online survey platform* mit Unternehmenssitz in Toronto, Kanada. Die Plattform besitzt eine große Nutzer-Community, da die Plattform auch über Social-Media-Funktionen wie Direktnachrichten verfügt. Plattform-Nutzern ist es nicht nur möglich Umfragen auszufüllen, sondern auch eigene Umfragen einzustellen und dadurch Punkte zu erlangen. Anders als bei vielen Crowdsourcing-Plattformen erfolgt die Vergütung nicht durch direkte finanzielle Entlohnung, sondern durch ein Punktesystem, wobei die Punkte dazu genutzt werden können, um Gutscheine und Geschenkkarten (z. B. bei Amazon oder Starbucks) zu erhalten. Die meisten Umfragen sind sehr kurz und dauern nur wenigen Minuten. Unternehmen oder Forschungseinrichtungen können die auf der Plattform registrierten Teilnehmer auch als eine Art Online-Panel nutzen. Ausgewählte

---

<sup>8</sup><http://www.tellwut.com>

Umfragen können nur an bestimmte Nutzer versendet werden. Auf der Produktseite von Tellwut ist dies als *sample solutions* beworben.<sup>9</sup> Im vorliegenden Fall wurde der Fragebogen nur an Tellwut-Nutzer versandt, die auch über einen Twitter-Account verfügen. Zur Registrierung auf der Plattform sind nur Staatsbürger der USA und Kanada berechtigt, die das 18. Lebensjahr vollendet haben. Unter Nutzern verfügt die Umfrage-Plattform über ein relativ hohes Renommee. Auf

The screenshot shows the 'United States Online Surveys' page on SurveyPolice.com. It features a table of top-rated online surveys. Tellwut is listed at rank 4 with a 5-star rating, 76/100 score, and 187 reviews. A 'JOIN NOW' button is visible next to each entry.

Rank	Survey Platform	Rating	Score	Reviews	Action
1	Pinecone Research	★★★★★	86 / 100	483 reviews	Join Now
2	Darwin's Data	★★★★★	84 / 100	318 reviews	Join Now
3	PaidViewpoint	★★★★★	78 / 100	707 reviews	Join Now
4	Tellwut	★★★★★	76 / 100	187 reviews	Join Now
5	MyView	★★★★★	67 / 100	192 reviews	Join Now
6	BzzAgent	★★★★★	67 / 100	116 reviews	Join Now
7	Opinion Outpost	★★★★★	66 / 100	1066 reviews	Join Now
8	Epoll Surveys	★★★★★	66 / 100	128 reviews	Join Now
9	OneOpinion	★★★★★	66 / 100	37 reviews	Join Now
10	YouGov	★★★★★	66 / 100	107 reviews	Join Now

**Abbildung 5.5.:** Das Ranking amerikanischer Umfrage-Plattformen auf *SurveyPolice.com*. Tellwut belegt dort im November 2016 Rang 4.

(Screenshot von: <https://www.surveypolice.com/countries/united-states>)

der Website *SurveyPolice.com*<sup>10</sup>, welche online-basierte Umfrage-Plattformen auf Glaubwürdigkeit und Vertrauen bewertet, belegt Tellwut in der Reihe der amerikanischen Umfrage-Plattformen den vierten Rang (vgl. Abbildung 5.5). Tellwut wurde zudem bereits von anderen Forscherinnen und Forschern aus unterschiedlichen Disziplinen verwendet, was auch aus der Perspektive des Umfrage-Einstellers für die Nutzung dieser Umfrage-Plattform spricht (Teevan et al., 2013; Levordashka & Utz, 2016). Für Tellwut musste der Fragebogen noch einmal erneut in dem von der Plattform zu Verfügung gestellten HTML-Formular umgesetzt werden. Abbildung 5.6 zeigt einen Ausschnitt des Fragebogens, wie er schließlich an das Tellwut-Panel gesendet wurde.

## 5.4. Methodendiskussion: Qualitative Auswertungsmethoden im Vergleich

Abschnitt 5.2 hat gezeigt, dass der Fragebogen eine Vielzahl an offenen Freitextfragen enthält. Im Folgenden wird hierzu eine Methodendiskussion durchgeführt, um geeignete Auswertungsmethoden für diese Freitextfragen zu eruieren.

<sup>9</sup><https://www.tellwut.com/online-survey-solutions>

<sup>10</sup><http://www.surveypolice.com>

Online Surveys, opinion polls, questions, earn rewards - Tellwut

<https://www.tellwut.com/survey/preview/id/46395>

Multiple Question Survey

Preview::

**Twitter Survey**

Dear User, I'm a research assistant at the chair for Information Science at the University of Regensburg. The purpose of this survey is to analyze Information Behaviour on Twitter. The questionnaire is voluntary and the data collected is strictly confidential. All participants will NOT be identified. All data and measurements obtained from this research study will be stored confidentially. Only researcher will have access to view any data collected during this research. This survey should only take about 7-10 minutes of your time. Questions to specific aspects of your Twitter account will occur. So for your convenience open a separate browser window and log in to Twitter.

1.1 I identify my gender as ...

☐ Female

☐ Male

☐ I would rather not say

1.2 When were you born?

☐ After 1994

☐ 1985-1994

☐ 1975-1984

☐ 1965-1974

☐ Before 1965

**Abbildung 5.6.:** Ausschnitt der Umfrage für die Tellwut-Probanden

Grundsätzlich stellt sich die Frage, wie das Spektrum der qualitativen Forschungs-/ Auswertungsmethoden aussieht und welche qualitativen Auswertungsmethode im Kontext qualitativ-orientierter Studien zum Informationssuchverhalten Verwendung finden. Wichtig ist zu erwähnen, dass es sich um Methoden der Datenauswertung und nicht der Datenerhebung handelt. Qualitative Forschungsmethoden finden sich vor allem im Methodeninventar der Sozial- und Gesellschaftswissenschaften und firmieren vor allem unter dem Begriff der „Methoden empirischer Sozialforschung“ (Strübing, 2013; Lamnek, 2010). Im angelsächsischen Raum liegt auch ein besonderer Fokus auf qualitativen Methoden im medizinischen Umfeld der Krankenpflege, sog. *nursing studies*<sup>11</sup> (Elo & Kyngäs, 2008; Wollny & Marx, 2009).

In den Methodenkapiteln der Hand- und Lehrbücher zu den Sozial- und Gesellschaftswissenschaften findet sich eine beinahe unüberschaubare Vielzahl „an Methoden zur interpretativen und rekonstruktiven Verarbeitung qualitativer Daten.“ (Strübing, 2013, S.131). Renata Tesch nimmt bereits 1990 eine Charakterisierung von 30 qualitativen Forschungsmethoden vor, welche am Forschungsinteresse der jeweiligen Methoden ausgerichtet ist und in Abbildung 5.7 aufgeführt (Tesch, 1990, S.72) sind.

Konzentriert man sich auf Methoden, die auf die Datenauswertung fokussieren, wird das Methodenspektrum etwas reduziert. Strübing führt beispielsweise folgende Auswertungsmethoden an (Strübing, 2013):

- *Grounded Theory*
- Objektive Hermeneutik
- Dokumentarische Methode
- Narrationsanalyse und Bibliographieforschung
- Ethnomethodologische Konversationsanalyse
- Diskursanalyse

<sup>11</sup>Auch die *Grounded Theory*, welche im Folgenden relevant werden wird, ist in diesem Umfeld entstanden.

THE CHARACTERISTICS OF LANGUAGE	THE DISCOVERY OF REGULARITIES	THE COMPREHENSION OF THE MEANING OF TEXT OR ACTION	REFLECTION
content analysis discourse analysis ethnography of communication ethnoscience structural ethnography symbolic interactionism ethnomethodology	transcendental realism ethnographic content analysis event structure analysis ecological psychology grounded theory phenomenography qualitative evaluation action research emancipatory/critical research holistic ethnography educational ethnography naturalistic inquiry	phenomenological research existential-phenom. Psychology empirical phenomenology dialogal phenomenology experiential phenomenology imaginal phenomenology case study life history hermeneutical research	educational Connoisseurship reflective phenomenology heuristic research

**Abbildung 5.7.:** Charakterisierung von qualitativen Forschungsmethoden nach Tesch (Tesch, 1990)

Weitere qualitative Auswertungsmethoden finden sich bei Mayring oder Kuckartz (Mayring, 2010; Kuckartz, 2012). Beide stellen vor allem die qualitative Inhaltsanalyse bzw. unterschiedliche Varianten der Inhaltsanalyse als Auswertungsmethoden für qualitative Daten vor. Nach Schreier ergeben sich für die praktische Anwendung qualitativer Methoden zwei Probleme, die einerseits auf „Unklarheiten und Unsicherheiten bei der Begriffsverwendung zurückzuführen sind“ (Schreier, 2014, S.2) — so zählt Krippendorff u.a. die Diskurs- und die Konversationsanalyse zu den qualitativ-inhaltsanalytischen Techniken — sowie andererseits „auf Hinweise und Ausführungen zu verschiedenen Varianten eines Verfahrens“ (Schreier, 2014, S.3) in der Literatur zurückgehen. Am stärksten ist dies bei der Qualitativen Inhaltsanalyse der Fall, bei der sich mindestens elf Varianten unterscheiden lassen (Schreier, 2014, S.5). Neben der Vielzahl an unterschiedlichen Auswertungsmethoden und deren Bezeichnungen, die Verwirrung stiften, ergeben sich noch weitere Probleme, die der Tatsache geschuldet sind, dass viele Methoden eine unterschiedliche sozial- und erkenntnistheoretische Rahmung besitzen, wodurch eine detaillierter Abgleich mit der eigenen wissenschaftstheoretischen Grundposition und der durch die Methode vorgegebenen erkenntnistheoretischen Basis durchgeführt werden sollte (Strübing, 2013).

Aufgrund der interdisziplinären Ausrichtung der Informationswissenschaft und ihrer Nähe zu den sozial- und gesellschaftswissenschaftlichen Disziplinen, welche besonders zu Beginn der 1980er Jahre wächst, ist es nicht verwunderlich, dass zahlreiche qualitative Forschungsmethoden aus diesem Bereich auch in das Methodeninventar der Informationswissenschaft Einzug halten. Dies ist vor allem der Tatsache geschuldet, dass viele ForscherInnen der Meinung sind, dass vor allem qualitative Methoden dazu geeignet sind, um die Bedürfnisse aufzudecken, die Informationssuchverhalten zu Grunde liegen und es somit ermöglichen, ein gründliches Bild des Nutzers und dessen Motive zu erhalten (Wilson, 1981, S.11). Eine Methode, der besondere Aufmerksamkeit zukam, ist die von Glaser und Strauss begründete *Grounded Theory* (Ellis, 2011, S.25). Wesentlichen Einfluss darauf, dass der Forschungsstil der *Grounded Theory* bei Studien zum Informationssuchverhalten eingesetzt wurde, hatte Wilson. Dieser betreute an der Universität Sheffield zahlreiche Qualifikationsarbeiten<sup>12</sup> und wendete diese Methode zum ersten Mal auch auf informationswissenschaftliche Fragestellungen an. Wie die Methodenetablierung in der Informationswissenschaft und der Überblick über weitere qualitative Auswertungsmethoden zeigt, gibt es zwei Methoden, denen besondere Aufmerksamkeit zukommt: die *Grounded Theory* sowie die Qualitative Inhaltsanalyse. Beide werden näher vorgestellt, da die Auswertung der eigenen Daten mithilfe einer Methoden-Mischung dieser beiden Verfahren durchgeführt wurde.

<sup>12</sup>Unter anderem die Promotionsarbeit von David Ellis.

### 5.4.1. Die Methode *Grounded Theory*

Die von Barney Glaser und Anselm Strauss 1967 im Rahmen eines soziologischen Forschungsprojektes zu „Bewusstseinskontexten während des Sterbeprozesses im Krankenhaus“ etablierte *Grounded Theory* (im folgenden GT) ist weniger eine bloße Methode zur Auswertung qualitativer Daten als vielmehr eine Methodologie, ein eigener Forschungsstil (Strübing, 2013, S.112) (Strauss & Corbin, 1996) (Wollny & Marx, 2009, S.470). Dies wird insbesondere dadurch deutlich, dass zwischen Auswertungsverfahren und Forschungsprozessorganisation ein innerer Zusammenhang besteht, der nicht ohne weiteres aufgelöst werden sollte (Strübing, 2013, S.109). Die GT ist ein weitverbreitetes Verfahren, welches Elemente/Techniken beinhaltet, die auch in anderen Methoden (in leicht veränderter Form) Anwendung finden (Strübing, 2013, S.109). In der Literatur wird häufig auf die Entwicklung unterschiedlicher Strömungen der GT hingewiesen, die einem über die Jahre wachsenden Dissens zwischen Glaser und Strauss geschuldet sind (Breuer, 2009, S.111ff.). Im Folgenden wird vor allem die Strausssche-Schule der GT und deren Charakteristika vorgestellt. Die GT nach Glaser besitzt wesentliche Unterschiede in Bezug auf die Verwendung bestehender Theorien, dem vorhandenen Kontextwissen, dem Kodierprozess (Breuer, 2009, S.111ff) (Truschkat et al., 2005, S.5), sowie eine, laut Strübing, nicht mehr zeitgemäße erkenntnistheoretische Position (Strübing, 2013, S.111).

Wesentliches Ziel der Straussschen Ausrichtung der GT ist die Erstellung einer empirischen, auf den Daten beruhenden Theorie menschlichen Handelns, einer sog. *Grounded Theory*, wobei *grounded* auch mit gegenstands- oder datenverankert übersetzt werden kann (Strauss & Corbin, 1996, S.XIII). Theorie bedeutet in diesem Sinne, dass das Ergebnis einer GT weit über die inhaltliche Beschreibung der Daten hinaus geht, indem ähnliche Daten zu Konzepten zusammengefasst werden und diese Konzepte schließlich miteinander verknüpft werden, wodurch übergeordnete Kategorien entstehen (Strauss & Corbin, 1996, S.13f.).

Mit der GT sind einige forschersische Grundprinzipien oder auch Verfahrensvorschläge verbunden, welche sich von anderen Forschungsprozessen unterscheiden und den wesentlichen Charakter der GT darstellen. Wie für andere Forschungsmethoden üblich, gibt es in der GT keine theoretischen Vorüberlegungen. Somit wird vorab auch keine empirisch zu überprüfende Hypothese formuliert. Ein weiteres wesentliches Merkmal der GT ist die Zirkularität respektive Parallelität von Datenerhebung, Datenauswertung und Entwicklung der Theorie. Diese Parallelität hat zur Folge, dass schon während der Datenerhebung mit der Auswertung dieser begonnen wird. Strübing spricht hier von der GT als „einzelfallanalytisches Verfahren“, da bereits mit dem ersten Fall die Auswertung und schließlich auch die Erstellung der gegenstandsverankerten Theorie beginnt (Strübing, 2013, S.113). Weitere Charakteristika bzw. Bestandteile der GT sind:

- Theoretisches Sampling
- Theoretische Sensibilität
- Datenanalyse

Einige dieser Charakteristika werden im Folgenden näher vorgestellt.

#### **Theoretisches Sampling**

Als Resultat dieser Parallelität aus Datenerhebung und -auswertung bedarf es auch einer speziellen Samplingstrategie, also einer speziellen Auswahl und Zusammenstellung von Versuchsobjekten zur Stichprobe. Dieses Vorgehen wird in der GT *theoretisches Sampling* genannt (Corbin & Strauss, 2008, S.143ff.). Theoretisches Sampling bezeichnet die sukzessive Zusammenstellung der Stichprobe während des Projektverlaufes in Abhängigkeit zu den Resultaten der Datenauswertung und der sich zu schärfen beginnenden Theorie. Dies bedeutet, dass die Kriterien zur Stichprobenzusammenstellung auf Konzepten und Kategorien beruhen, die während der Datenauswertung entstehen.

Dies steht im Einklang mit dem Verhältnis von Theorie bzw. Vorwissen und der Formulierung der Hypothesen, denn eine kriterienorientierte Stichprobenauswahl würde mehr Vorwissen voraussetzen, welches bei der GT vermieden werden sollte (Strübing, 2013, S.117). Theoretisches Sampling wirft im Vergleich zur herkömmlichen Stichprobenkonstruktion zwei Fragen auf: einerseits die Frage nach der Wahl des sog. *ersten Falls*, da dieser nachhaltig auch die Auswahl weiterer Fälle beeinflusst und andererseits die Frage, wann man genügend Fälle für die Erstellung der Theorie ausgewählt hat und somit die sog. *theoretische Saturation* erreicht ist. Zweiteres beantworten Corbin und Strauss mit: „when no new data are emerging“ und auch trotz eines neuen Falles keine neuen Konzepte bzw. Veränderungen in der Kategorie zu erkennen sind (Corbin & Strauss, 2008, S.143). Ersteres kann nicht wirklich gelöst werden. Seine Bedeutung wird für die Bildung der Theorie in der Literatur aber oft marginalisiert (Strübing, 2013, S.117). In der Literatur wird auf die Problematik hingewiesen, dass diese Samplingstrategie aufgrund von forschungsökonomischen Aspekten (Zugang der ForscherIn zum Feld, Art der Datenerhebung, etc.) nie zu hundert Prozent durchgeführt werden kann (Strübing, 2013, S.117). Truschkat, Kaiser und Reinartz votieren deshalb für eine flexible Handhabung der Samplingstrategie (Truschkat et al., 2005, S.2). Ferner schildern die Autorinnen, dass theoretisches Sampling oft falsch interpretiert wird und nicht die bloße Erhöhung der Datenbasis das Ziel sein sollte, sondern, dass auch Vergleiche innerhalb des bestehenden Datenbestandes Teil des theoretischen Samplings sind:

„Es geht also beim theoretischen Sampling nicht unbedingt darum, die Datenbasis zu erhöhen, sondern es können auch aus der bestehenden Datenbasis heraus Vergleiche hergestellt werden. Wir entdecken systematisch auch am Material Kontraste und Ähnlichkeiten. D.h. wir fahren fort, theoretisch zu 'samplen'. Nur sind es jetzt nicht mehr Fälle oder Situationen, die wir aufeinander beziehen, sondern das erhobene Datenmaterial selbst“ (Alheit 1999, S. 16, zitiert nach Truschkat et al. (2005)).

### Theoretische Sensibilität

Darüber hinaus stellt die GT auch einige Anforderungen an den Forschenden selbst. Eine oft angesprochene Fähigkeit, die GT-Praktizierende besitzen sollten, ist die *theoretische Sensibilität* (Strauss & Corbin, 1996, S.25). Sie bezeichnet das Bewusstsein der ForscherIn „für die Feinheit in der Bedeutung von Daten“ (Strauss & Corbin, 1996, S.25). Diese ist von vielen unterschiedlichen Faktoren abhängig z. B. der gelesenen Literatur, der Allgemeinbildung der Person sowie persönlicher und beruflicher Erfahrung (Strauss & Corbin, 1996, S.25ff.). Strauss und Corbin schlagen aber auch Methoden vor, wie die Theoretische Sensibilität erhöht werden kann. Wollny und Marx fassen diese wie folgt zusammen (Wollny & Marx, 2009, S.471):

- Analyse durch Fragestellung (W-Fragen)
- Analyse eines Wortes/Satzes (Line-by-Line-Analyse)
- Analyse durch Vergleiche (Phänomen-Konzept-Abgrenzung)

Neben der Theoretischen Sensibilität wird von der ForscherIn auch ein gewisses Maß an Kreativität gefordert. Diese hilft dabei „Kategorien treffend zu bezeichnen, seine Gedanken schweifen zu lassen, freie Assoziationen zu bilden, die für das Stellen anregender Fragen notwendig sind, und Vergleiche anzustellen, die zu neuen Entdeckungen führen.“ (Strauss & Corbin, 1996, S.12). Mit diesem Argument sind Strauss und Corbin aber nicht alleine. Viele Wissenschaftler sind der Ansicht, dass qualitative Datenanalyse neben einem wissenschaftlichen auch immer einen künstlerischen Charakter besitzt (Truschkat et al., 2005).

### Datenanalyse

Herzstück der GT stellt die Datenanalyse in Form eines sukzessiven bzw. iterativen Kodierprozesses dar. Das Interpretieren und Analysieren des empirischen Materials geschieht dabei in drei Phasen des Kodierens, bei denen Konzepte bzw. Codes erarbeitet werden, „[...] die einzelnen Ereignissen, Vorkommnissen oder anderen Beispielen für Phänomenen zugeordnet werden“ (Strauss & Corbin, 1996, S.43). Diese Codes werden später zu Kategorien zusammengefasst bzw. zu einem Konzept



höherer Ordnung gruppiert. Strübing führt an, dass die Begriffe *Kode* und *Kategorie* in der Sekundärliteratur zur GT oft gleichbedeutend verwendet werden (Strübing, 2013, S.124). Die Phasen des Kodierens sind (Strauss & Corbin, 1996, S.43):

- Offenes Kodieren
- Axiales Kodieren
- Selektives Kodieren

#### **Offenes Kodieren**

Offenes Kodieren ist „[d]er Prozeß des Aufbrechens, Untersuchens, Vergleichens, Konzeptualisierens und Kategorisierens von Daten“ (Strauss & Corbin, 1996, S.43). Während der Phase des offenen Kodierens erfolgt eine intensive Auseinandersetzung mit dem Material. Strauss und Corbin schlagen hierfür die detaillierte Analyse jeder einzelnen Zeile, die sogenannte Line-by-Line-Analyse, vor. Dabei werden erste Konzepte, Ideen und Hypothesen formuliert und in Memos/Notizen festgehalten. Die Daten werden dabei systematisch aufgeschlüsselt und verdichtet (Wollny & Marx, 2009, S.471). Wichtig dabei ist, dass während des offenen Kodierens die theoretische Sensibilität der ForscherIn geschärft wird, wobei oben aufgeführte Techniken/Methoden zum Einsatz kommen. Beim offenen Kodieren gilt es, eine optimale Distanz zwischen Daten und gebildeten Kodes zu finden, was eine Frage der optimalen Korngröße also einer makroskopischen bzw. mikroskopischen Auseinandersetzung mit den Daten betrifft (Breuer, 2009).

#### **Axiales Kodieren**

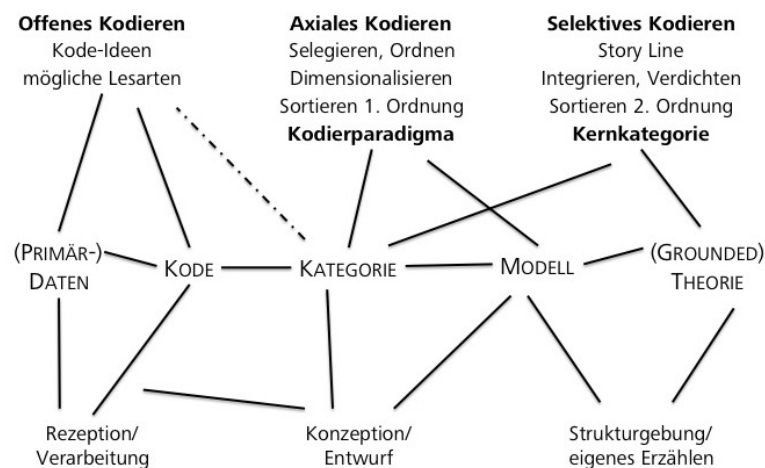
Die zweite Phase des Codierens definieren Strauss und Corbin als „[e]ine Reihe von Verfahren mit denen durch das Erstellen von Verbindungen zwischen Kategorien die Daten nach dem offenen Kodieren auf neue Art zusammengesetzt werden“ (Strauss & Corbin, 1996, S.75). Um die Verbindung von Kategorien/Subkategorien herausarbeiten zu können, bedarf es wieder einiger Techniken, wie der des ständigen Vergleichs. Führt der ständige Vergleich im Schritt des offenen Kodierens zu Konzepten, wird beim axialen Kodieren schließlich versucht, durch Hinzunahme bestimmter Faktoren wie Kontext, Strategie, Phänomen und Handlung, konkrete Bezeichnungen für die Abgrenzungsfaktoren zwischen den Konzepten zu finden und diese für die Erstellung von Kategorien, welche als Gruppen aus Konzepten geschaffen werden, zu nutzen (Strauss & Corbin, 1996, S.74ff.). Für jedes Konzept wird nach Ähnlichkeiten, Gegenteiligen und Zusammenhängen zu anderen Konzepten gesucht (Breuer, 2009, S.77). Das axiale Kodieren widmet sich folglich noch den „einzelnen empirischen Vorkommnissen“ (Strübing, 2013, S.121) und versucht diese detailliert zu ergründen. Die Beantwortung der allumfassenden Forschungsfrage steht hier noch im Hintergrund.

#### **Selektives Kodieren**

Die Phase des selektiven Kodierens wird von Strauss und Corbin als „[d]er Prozeß des Auswählens der Kernkategorie, des systematischen In-Beziehung-Setzens der Kernkategorie mit anderen Kategorien“ bezeichnet (Strauss & Corbin, 1996, S.94). Ziel ist das Finden einer Kernkategorie, welche andere Kategorien integriert. Dies gelingt nach Strauss und Corbin mit der Technik des *roten Fadens der Geschichte*. Als Schlüsselkategorie sollte die Kategorie ausgewählt werden, mit der die Forschende das Forschungsprojekt am besten gelöst sieht.<sup>13</sup> Tatsächlich führen Strauss und Corbin noch weitere Kriterien für die Kernkategorie an (Strauss & Corbin, 1990, S.67). Breuer bezeichnet das selektive Kodieren auch als axiales Kodieren auf einem „Niveau höherer Abstraktion, allgemeinerer Theoretisierungsstufe“ (Breuer, 2009, S.92). Dies entspricht der Maxime des selektiven Kodierens, insofern daraus schließlich die GT als Theorie entsteht. Der Übergang zwischen den Phasen ist dabei fließend, iterativ und kollaborativ und es findet ein ständiger Wechsel innerhalb der Phasen statt. Generell kommt dem Forschen im Team und der Diskussion von analytischen Zwischenschritten in der GT ein hohes Maß an Bedeutung zu. Dies ist auch ein Zeichen

<sup>13</sup>Es kommt immer wieder die Frage auf, ob es exakt eine Kernkategorie sein muss, oder ob es auch mehrere Kategorien sein können.

von Qualitätssicherung und Güteprüfung. „Gerade das offene und das axiale Kodieren (s.u.) sind Arbeitsschritte, in denen mehrere Forscherinnen einander ideal ergänzen, aber auch wechselseitig in der Spur halten können“ (Strübing, 2013, S.114). Kollaboration und Diskussion zwischen mehreren Beteiligten hilft also dabei, die Interpretation von Textstellen und die Generierung von Codes im interpretativen Rahmen zu halten (Strübing, 2013, S.114). Strübing setzt sich detailliert mit der Frage der Güteprüfung und Qualität von GT auseinander. Er kommt dabei zu dem Schluss, dass die klassische Trias der Gütekriterien (Reliabilität, Repräsentativität und Validität) für die Qualität von empirischer Arbeit auch auf GT anwendbar sind (Strübing, 2008, S.80). Darüber hinaus stellt er fest, dass die Charakteristika der GT, wie theoretisches Sampling, Schreiben von Memos etc., inhärente qualitätssichernde Maßnahmen sind (Strübing, 2008, S.80). Ihre wahre Qualität stellt eine GT aber in der praktischen Angemessenheit unter Beweis, getreu dem Motto „Praxis bringt die Validierung der Theorie“ soll eine auf Daten fundierte Theorie auch von praktischer Relevanz sein und Handlungsfähigkeit ermöglichen (Strübing, 2008, S.85). Dieser Sachverhalt wird in Abschnitt 5.6.3.4 noch näher erläutert und als Ziel der vorliegenden Analyse formuliert.



**Abbildung 5.8.:** Kodierprozess der *Grounded Theory* in systematischer Anordnung nach Breuer (Breuer, 2009)

Eine Visualisierung des Kodierprozesses erstellt Breuer (vgl. Abbildung 5.8). Betrachtet man die Darstellung des Kodierprozesses zeilenweise, finden sich in der ersten Zeile die unterschiedlichen Kodierphasen, in der zweiten Zeile, das Material auf welche sich diese Phasen beziehen und in der dritten Zeile die Zielsetzungen und Konstruktionslogiken der Phasen (Breuer, 2009, S.76). Dass es sich hier um Linien handelt, die die einzelnen Bestandteile nur lose miteinander verknüpfen, liegt an der nicht stringent vorgegebenen Richtung und versucht der Tatsache Rechnung zu tragen, dass bestimmte Tätigkeiten übersprungen bzw. auch mehrmals iterativ wiederholt werden können.

Trotz vieler Vorgaben ist die GT nicht so restriktiv, wie diese Charakteristika auf den ersten Blick suggerieren. Die Vorgaben sind mehr als „Leitlinien zu verstehen und nicht als starre, zwingende Handlungsanweisungen“ ohne Spielraum für die ForscherIn (Strübing, 2013, S.117). So ist die GT nicht auf eine bestimmte Datenerhebungsmethodik angewiesen und von Freitextfragen, Interviewtranskripten und Chatprotokollen über Bilder und Videos auch auf selbsterstellte ethnographische Beobachtungsprotokolle oder sogar quantitative Daten anwendbar (Strübing, 2013, S.126).

In diesem Kontext argumentieren Truschkat, Kaiser und Reinartz, man müsse sich auch vor Augen führen, dass die Anleitung zur Durchführung der GT immer auch eine idealtypische Form der Forschung darstellt, die nie in jedem Fall exakt eingehalten werden kann (Truschkat et al., 2005, S.11). Vor allem bei der Auswertung der Daten mithilfe der Kodierverfahren ist die GT variabel und anpassungsfähig.

### 5.4.2. Die Methode qualitative Inhaltsanalyse

Eine ganzheitliche Auseinandersetzung mit der Methode der qualitativen Inhaltsanalyse (qIA) nimmt Margrit Schreier vor. Sie argumentiert, dass „[...]kein Konsens darüber [besteht], was qualitative Inhaltsanalyse ausmacht“ (Schreier, 2014, S.2). Zunächst besitzt die Methode unterschiedliche Traditionen in der deutschsprachigen und englischsprachigen Literatur, wobei Publikationen in letzterer eher der quantitativen Schule zuzuordnen sind (Schreier, 2014, S.1).<sup>14</sup> Im deutschsprachigen Raum haben sich vor allem die Werke von Mayring und Kuckartz über qualitativ-inhaltsanalytische Verfahren etabliert (Mayring, 2010; Kuckartz, 2012; Mayring, 2014). Beide verankern die qIA in der Hermeneutik, betonen aber, dass die Qualitative Inhaltsanalyse stärker regelgeleitet ist und wesentliche Gütekriterien besitzt (Schreier, 2012, S.1) (Kuckartz, 2012, S.31ff.). Schreier definiert die qualitative Inhaltsanalyse als „ein Verfahren zur Beschreibung ausgewählter Textbedeutungen“ (Schreier, 2014, S.1), wobei die Beschreibungen dieser Bedeutungen in Form von Kategorien innerhalb eines Kategoriensystems manifestiert werden. Die Kategorien des Kategoriensystems werden dabei induktiv erstellt, was bedeutet, dass dieses dazu in der Lage ist, die wesentlichen Bedeutungsaspekte der Daten bzw. des Materials zu erfassen. Nach Schreier besitzt die qualitative Inhaltsanalyse folgende wesentliche Charakteristika:

- Kategorieorientierung
- Interpretatives Vorgehen
- Einbeziehung latenter Bedeutungen
- Entwicklung eines Teils der Kategorien am Material
- Systematisches, regelgeleitetes Vorgehen
- Orientierung an Reliabilität und Validität

Wie bereits erwähnt, stellt die Vielzahl der in der Literatur beschriebenen Varianten der qualitativen Inhaltsanalyse die Forschenden in der Praxis vor Probleme. Beispielsweise unterscheidet Mayring drei Hauptformen der Inhaltsanalyse, wobei er die strukturierende qIA zusätzlich in die inhaltliche, die typisierende und die skalierende Strukturierung unterteilt (Mayring, 2010, S.48ff.). Kuckartz unterscheidet neben der inhaltlich-strukturierenden qIA die evaluative, sowie die typenbildende qIA (Kuckartz, 2012). In der englischsprachigen Literatur unterscheiden Hsieh und Shannon zwischen der *conventional*, der *directed* und der *summative* Variante der qIA. Elo und Kyngäs unterscheiden zwischen *inductive* und *deductive content analysis* (Elo & Kyngäs, 2008). Der ForscherIn stellt sich somit die Frage: Welche ist die geeignetere Variante? Wie verhalten sich die Varianten zueinander? Welche Unterschiede und Gemeinsamkeiten besitzen sie? Mit einem 2014 erschienen Artikel gibt Margrit Schreier Antworten auf diese Fragen und bietet in Form eines Werkzeugkastenmodells Orientierung für auf Hilfe angewiesene Anwender der qIA (Schreier, 2014).

Trotz der unterschiedlichen Varianten der qIA identifiziert Schreier die inhaltlich-strukturierende Inhaltsanalyse nach Mayring, sowie die Inhaltsanalyse durch Extraktion nach Gläser und Laudel

<sup>14</sup>Kuckartz führt hierzu an, dass der englische Begriff der *content analysis* auch häufig ausschließlich auf eine automatische, frequenzstatistische Analyse abzielt, welche im Deutschen als computerunterstützte Inhaltsanalyse (CUI) bezeichnet wird (Kuckartz, 2012, S.39).

(Gläser & Laudel, 2009), als die beiden Kern- bzw. Basisverfahren. Als charakteristisch für die strukturierende Form der qIA nennt Schreier folgende Punkte (Schreier, 2014):

- Iteratives, induktives Vorgehen
- Erstellung eines Kategoriensystems
- Probekodierung und sukzessive Weiterentwicklung des Kategoriensystems
- Anwendung des Kategoriensystems auf das gesamte Datenmaterial

Von diesem Verfahren der qIA grenzt sich die qIA durch Extraktion deutlich ab. Sie ist aktueller und besitzt folgende Charakteristika:

- Zusammenfassung von Verdichtung als Methoden der Kategoriengenese
- stark theoriegeleitet
- Entstehung eines Paralleltextes
- Keine Probe- oder Doppelkodierung

Schreier führt weiter aus:

„Für alle übrigen Varianten, die in der Literatur genannt werden, gilt jedoch, dass sie nicht notwendig als eigenständige Verfahren gelten können, sondern an einer bestimmten Stelle von der generischen Variante qualitativ-strukturierende Inhaltsanalyse abweichen, bzw. diese in einer bestimmten Weise spezifizieren“ (Schreier, 2014, S.22)

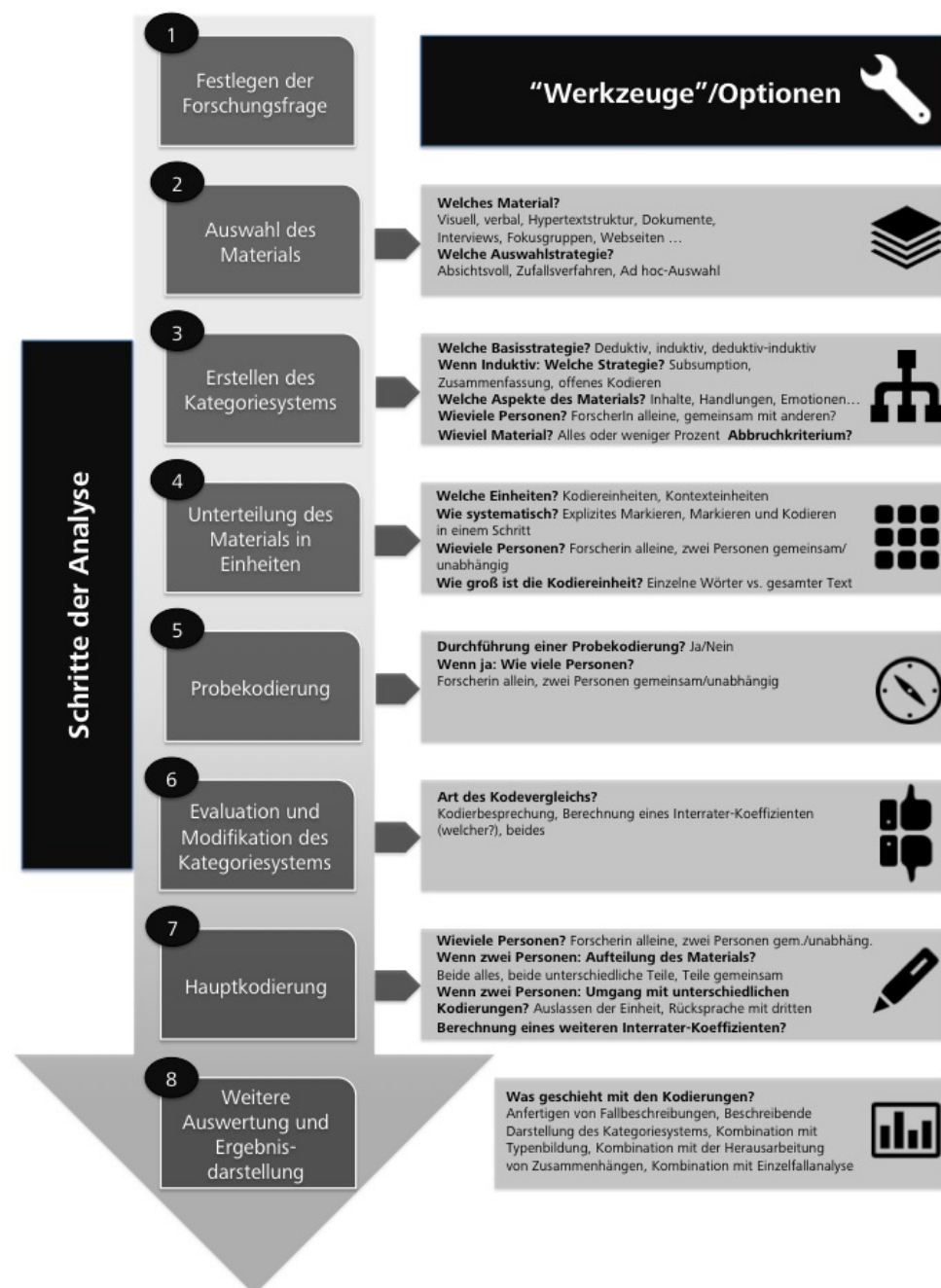
Es wird in der Literatur eine willkürliche und keineswegs erschöpfende Aufzählung von Varianten vorgenommen.<sup>15</sup> Diese Tatsache motiviert Schreier zur Erstellung eines qualitativ-inhaltsanalytischen Werkzeugkastenmodells — wobei die inhaltlich-strukturierende qIA als Basisablauf dient — welcher zu bestimmten Analyseschritten Werkzeuge oder Optionen bietet, zwischen denen eine Auswahl getroffen werden kann, um das Verfahren der Inhaltsanalyse entsprechend der eigenen Bedürfnisse zu adaptieren.

Abbildung 5.9 zeigt das Werkzeugkastenmodell. Insgesamt unterteilt Schreier die qIA in acht Analyseschritte. Zu jedem dieser Schritte bietet das Modell Werkzeuge oder Optionen, die in Form von Fragen formuliert sind und der ForscherIn als Orientierung bzw. zur Entscheidungsfindung dienen. In Analyseschritt 2 *Auswahl des Materials* wird in Form der Frage *Welches Material?* eine Möglichkeit von Artefakten aufgezählt, die analytisch analysiert werden können und somit den Raum der möglichen Optionen bzw. unterschiedlicher inhaltsanalytischer Verfahren aufzeigen. Je nachdem welche Optionen bzw. Werkzeuge gewählt werden, ergeben sich eigene Spezifikationen der qualitativen Inhaltsanalyse. Die inhaltlich-strukturierte Qualitative Inhaltsanalyse orientiert am Werkzeugkasten, bietet also eine Reihe vorab spezifizierter Schritte, welche zu vielfältigen Realisierungsmöglichkeiten der qIA führen. Diese Variabilität der qIA als Auswertungsmethode wurde auch für die Auswertung der eigenen Daten genutzt.

### 5.4.3. Beschreibung des Vorgehens bei der Datenauswertung

Die Beschreibung der eigenen Datenanalyse beginnt mit der Argumentation für die Wahl einer Methodenmischung. Aus Verletzungen bestimmter methodischer Grundprinzipien der GT mussten Konsequenzen gezogen werden. Da die Datenauswertung nicht durch eine reine *Grounded Theory Methodology* geleitet durchgeführt werden konnte, wird eine Kombination mit der qIA durchgeführt. Es wurde bereits angesprochen, dass die methodologischen Vorgaben der GT nicht so restriktiv

<sup>15</sup>Alleine bei der Erstellung von Kodes und Konzepten könnte man nach Schreier zwischen induktiv, deduktiv und gemischt induktiv-deduktiv unterscheiden, — eine Differenzierung die auch von Hsieh und Shannon getroffen wird — oder zwischen der Anzahl der Kodierer und Kodierschritte, woraus sich aus verschiedenen Konstellationen verschiedene Varianten der Inhaltsanalyse ergeben.



**Abbildung 5.9.:** Visualisierung des „Werkzeugkastenmodells der Inhaltsanalyse“. Eigene Darstellung nach Schreier (Schreier, 2014)

tiv sind wie sie erscheinen mögen. Auch Strübing formuliert dies so: „Die Forschenden [sollten] das vorgeschlagene Verfahren situativ an Gegenstand, Forschungsfrage und eigene Kompetenzen anpassen können“ (Strübing, 2014, S.92). Mit der Anpassung an den jeweiligen forschungsökonomischen und situativen Kontext vertritt Strübing eine Position, die von vielen anderen Forscherinnen und Forschern geteilt wird (Truschkat et al., 2005; Legewie, 2006). Obwohl eine Adaption der GT möglich ist und in der Praxis auch umgesetzt wird, gibt es einige wesentliche Elemente der GT,

die nicht verletzt werden dürfen.

„Im Lichte der in diesem Buch diskutierten methodologischen Argumente sind eine iterativ-zyklische Vorgehensweise, das theoretical sampling mit dem Kriterium de[r] theoretischen Sättigung und nicht zuletzt die Methode des ständigen Vergleichens unter Verwendung generativer Fragen ebenso unverzichtbar wie Kodieren und Memos schreiben.“ (Strübing, 2014, S.93).

Da für die Auswertung der Freitextfragen wesentliche Charakteristika der GT verletzt werden, wird eine Methodenkombination vorgeschlagen, was in der Praxis keine unübliche Vorgehensweise darstellt. So findet man bei Schreier das offene Kodieren der GT als eine Strategie zur datengetriebenen Erstellung von Kategoriensystem bzw. Kode-Schemata in der induktiven Vorgehensweise der qualitativen Inhaltsanalyse. „A final data-driven strategy is to adapt the steps of data analysis in grounded theory to generating inductive categories“ (Schreier, 2012, S.88). Häufig werden auch unterschiedliche Varianten der qualitativen Inhaltsanalyse miteinander kombiniert, wofür auch Margrit Schreiers Werkzeugkastenmodell ein Plädoyer darstellt (Schreier, 2014). Auch Heiner Legewie argumentiert, dass es sinnvoll sein kann „einzelne heuristische oder methodische Elemente der Grounded Theory herauszugreifen und mit anderen Methoden zu kombinieren“ (Legewie, 2006, S.10). Es werden nun kurz die wesentlichen Verstöße gegen die Forschungsprinzipien der GT genannt.

#### **Verstöße gegen Richtlinien der GT und Hinzunahme von Prinzipien der qIA**

Zwei wesentliche Charakteristika der GT werden bereits vor Auswertungsbeginn verletzt, da die Auswertung der Daten nicht ohne eine gewisse Theorie bzw. Vorannahmen über die Gründe für das Favorisieren von Tweets vonstatten ging und auch keine Distanzierung von der Forschungsfrage erfolgte. Sowohl Wollny und Marx als auch Truschkat, Kaiser und Reinartz sehen beide Faktoren als wichtig für die Durchführung einer GT (Wollny & Marx, 2009, S.473) (Truschkat et al., 2005). Wie man auch anhand des Werkzeugkastenmodells erkennen kann, geht die qIA flexibler mit den Fragen nach Theoriegeleitetheit, Vorwissen und Forschungsfrage um und ist sehr pragmatisch orientiert.<sup>16</sup> Da alle Daten vorab erhoben wurden, werden in der vorliegenden Studie Verletzungen gegenüber der Zirkularität von Datenauswertung und Datenerhebung vorgenommen. Somit wurde auch das Theoretische Sampling verletzt, das eine Zusammenstellung der Stichprobe auf der Basis der ersten Datenauswertung vorsieht. Ob dies eine schwerwiegende Verletzung darstellt ist diskussionswürdig, betrachtet man die Position von Truschkat, Kaiser und Reinartz in Abschnitt 5.4.1. Trotz dieser Verletzungen wurde versucht, die Vorgaben der GT einzuhalten. So wurde der iterative Kodierprozess der GT, welcher die drei Phasen des offenen, axialen und selektiven Kodierens vorsieht, vollständig in die eigene Auswertung integriert. Dieser wurde, wie von der GT erwartet, stets kollaborativ durchgeführt. An der GT als Rahmenmethode für die Auswertung wurde insbesondere deshalb festgehalten, weil das Ziel der Analyse eine gegenstandsverankerte Theorie nach der GT sein sollte, welche „praktisches Erklärungspotential“ (Strübing, 2013, S.127) bietet und „für Akteure im untersuchten Handlungsfeld rezipierbar bessere Verständnis ihrer Praxis [liefert]“ (Strübing, 2013, S.127). Ziel der Analyse der Favorisierungsgründe ist es, aus diesen Ergebnissen praktisches Handwerkszeug resultieren zu lassen, um zum Beispiel Features für Filter-Algorithmen entwickeln zu können. Darüber hinaus sollen Personen, die sich über die Möglichkeiten der Favorisierungsgründe informieren, indem sie diesen Text lesen, dazu motiviert sein, über ihr eigenes Handeln, also das Setzen von Favoriten, zu reflektieren und sich somit wieder zu erkennen, oder neue Möglichkeiten zu entdecken, wozu sie die Favoriten-Funktion nutzen können.

Ein weiterer Grund für die Hinzunahme von Prinzipien der qIA zum Auswertungsprozess war der Wunsch nach prüfbaren Gütekriterien wie der Prüfung der Reliabilität der Ergebnisse. In der qIA wird hierfür die aus der quantitativen Inhaltsanalyse bekannte Vorgehensweise der Berechnung

<sup>16</sup>Der Forschungsprozess der *Grounded Theory* erwartet zu Beginn der Datenerhebung und -auswertung keine konkrete Forschungsfrage. Dass das Werkzeugkastenmodell pragmatisch ausgerichtet ist, lässt sich auch daran erkennen, dass die Frage nach Theorie und Vorwissen gar nicht aufkommt.

einer Inter-rater-Übereinstimmung verwendet, um die Güte der Ergebnisse durch eine Maßzahl zu bestätigen und eine große Varianz im Kategoriensystem zu vermeiden (Wollny & Marx, 2009, S.468). Zur Berechnung der Inter-rater-Übereinstimmung wurde in beiden Fällen eine von Harris und Burke entwickelte Herangehensweise in Betracht gezogen, die die Berechnung einer Kappa-Variante auf der Basis der Intra-Klassen-Korrelation (ICC) vorsieht und somit ein *Multi Code Multi Coder Kappa* (MCMC) darstellt (Harris & Burke, 2005). Wie die Formel zeigt, ersetzt die Intraklassenkorrelation (ICC) das  $P$  bei der Berechnung des Kappa.

$$\kappa = \frac{p_0 - p_c}{1 - p_c} \Rightarrow \kappa = \frac{ICC_O - ICC_c}{1 - ICC_c}$$

Harris und Burke argumentieren, dass der ICC-Koeffizient als Reliabilitätsmaß dem Übereinstimmungsmaß Cohens  $\kappa$  dann vorzuziehen ist, wenn mehrere Kodierende nur jeweils einen Kode an ein Objekt vergeben. Da bei Harris und Burke — wie auch im vorliegenden Fall — mehrere Codes an ein Objekt vergeben werden können sollen, musste eine Adaption des  $\kappa$ -Maßes entsprechend ihren Überlegungen erfolgen (Harris & Burke, 2005). Es existieren jedoch auch Kappa-ähnliche Varianten, bei denen der Mittelwert aus den paarweisen Kappa-Werten berechnet werden kann. Dieses Vorgehen ist zum Beispiel bei Hallgren beschrieben (Hallgren, 2012, S.30). Das Skalenniveau stellt ebenfalls einen problematischen Parameter dar. Betrachtet man Kategoriensysteme, wie sie üblicherweise aus Verfahren der Inhaltsanalyse oder der GT resultieren, so sind die Ergebnisse des Kodierprozesses, in den seltensten Fällen intervallskalierte Daten. Dieses Skalenniveau wäre aber die Voraussetzung für die Berechnung der Intra-Klassen-Korrelation (Wirtz & Caspar, 2002, S.40ff.). Nach Wirtz und Caspar können Methoden zur Berechnung der Urteilsübereinstimmung, welche bei dichotomen Daten verwendet werden können, auch bei „nicht nominalskalierten Zeichensystemen [...] bei denen jedes Objekt mehreren Kategorien gleichzeitig angehören kann [...]“ (Wirtz & Caspar, 2002, S.47), genutzt werden. Bedingung hierfür ist, dass ein Umkodieren bzw. Dichotomisieren (vorhanden / nicht vorhanden) für jede einzelne Kategorie und jedes Objekt erfolgt (Wirtz & Caspar, 2002, S.47). Dichotome Daten wiederum können, so Wirtz und Caspar, „im Rahmen des allgemeinen linearen Modells wie metrische oder intervallskalierte Daten behandelt werden“ (Wirtz & Caspar, 2002, S.87). Das Vorgehen, wie es bei Harris und Burke beschrieben ist, ist somit auch in Bezug auf den Parameter des Skalenniveaus nachvollziehbar. Im Kapitel zur Analyse der Wiederfindensgründe werden zusätzlich die Werte für die Variante eines Kappa-Maßes für drei Kodierende — bei der der Mittelwert aus den paarweisen Kappa-Koeffizienten genommen wird — sowie das Ergebnis eines klassischen ICC als *baseline* bzw. Vergleichswerte berichtet.

## 5.5. Charakterisierung der Studienteilnehmer

Die Umfrage wurde von 606 Probanden ausgefüllt, welche ein *convenience sample* aus einem eigens gehosteten *snowball sample* und einem Crowdsourcing-Sample bilden. Die lokal gehostete, web-basierte Variante der Umfrage, welche als Snowball-sample-Verfahren via Mailverteiler, Foren und Social Media verbreitet wurde, resultierte in einer Stichprobe von 103 Teilnehmern. Das Crowdsourcing-Sample wurde über die Umfrageplattform TellWut.com rekrutiert, wobei 503 Teilnehmer zum Preis von \$ 200 angefordert wurden, was einer Belohnung von 40 Cent pro Teilnehmer für das Ausfüllen des Fragebogens entspricht.

Die Nutzung einer Crowdsourcing-Plattform zur Rekrutierung ermöglichte den Zugang zu einer größeren und zweifellos repräsentativeren Stichprobe an Twitter-Nutzern, als dies mit einem reinen *snowball sample* möglich gewesen wäre. Tabelle 5.1 gibt einen Überblick über die Geschlechter- und Altersverteilung der beiden Stichproben. Über die Hälfte der Teilnehmer (52,4%) aus dem

GESCHLECHT	Männlich	Weiblich	Ohne Angabe	Gesamt
Schneeball-Stichprobe	68 (66%)	34 (33%)	1(1%)	103
Tellwut-Stichprobe	135 (26,8%)	362 (72%)	6 (1,2%)	503
Gesamt	203 (33,5%)	396 (65,3%)	7(1,2%)	606

ALTER	<19	19-29	30-39	40-49	49>	Gesamt
Schneeball-Stichprobe	3 (2,9%)	54 (52,4%)	35 (34%)	4 (3,8%)	7 (6,8%)	103
Tellwut-Stichprobe	31 (6,6%)	171 (34%)	109 (21,7%)	85 (16,9%)	107 (21,3%)	503
Gesamt	34 (5,6%)	225 (37,1%)	144 (23,8%)	89 (14,7%)	114 (18,8%)	606

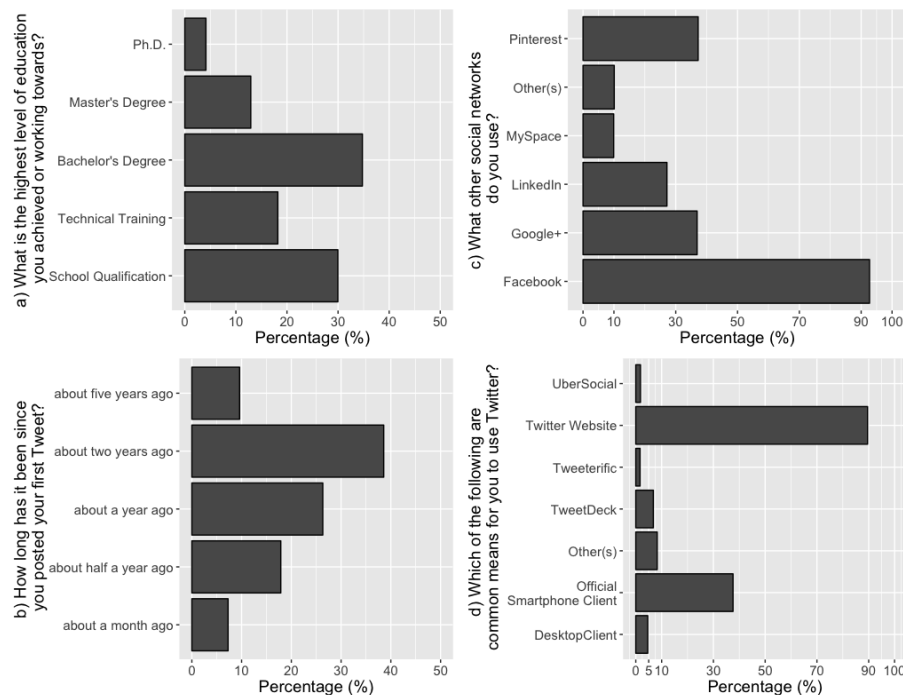
**Tabelle 5.1.:** Demographische Daten zu Alter und Geschlecht der beiden Stichproben im Vergleich

Schneeball-Sample sind zwischen 19-29 Jahren alt, was dem starken Verbreitungsfokus der Umfrage im universitären Kontext geschuldet ist. Das Crowdsourcing-Sample hingegen zeigt hier ein heterogenes Bild, da im Vergleich zum Schneeball-Verfahren mehr jüngere (6,6% der Teilnehmer sind jünger als 19 Jahre) und ältere Probanden (38,2% sind 40 Jahre alt oder älter) erreicht werden konnten. In Bezug auf die Geschlechterverteilung lassen sich ebenfalls schiefe Verteilungen feststellen. Im Schneeball-Sample ist die Verteilung von Männern zu Frauen 66% zu 33%, so dass hier ein deutlich höherer Anteil an männlichen Probanden vorhanden ist. Bei der Crowdsourcing-Stichprobe verkehrt sich dieser Trend ins Gegenteil. 72% der TeilnehmerInnen, die via Tellwut rekrutiert wurden, sind Frauen. Das Phänomen, dass via Crowdsourcing häufig mehr Frauen rekrutiert werden können, ist bekannt und in der Literatur bereits dokumentiert (Ross et al., 2010). Dort wird die Vermutung geäußert, dass es sich um sogenannte *stay at home mom's* handelt, die sich neben ihrer Tätigkeit als Hausfrau zusätzliches Geld durch bezahltes Crowdsourcing erarbeiten (Ross et al., 2010). Weitere demographische Charakteristika werden in Bezug auf die gesamte Stichprobe vorgestellt.

Die Abbildungen 5.10 a)-d) geben eine Zusammenfassung über weitere demografische Charakteristika. Grafik 5.10 a) zeigt das Bildungsniveau der Teilnehmer. Es wird ersichtlich, dass es sich um eine überwiegend gebildete Stichprobe handelt. Mehr als die Hälfte (51,9%) der Probanden hatte zum Zeitpunkt bereits einen Bachelor-Abschluss erworben oder diesen Ausbildungsabschnitt als nächstes Ziel erklärt. Abbildung 5.10 b) zeigt die Antworten auf die Frage *What other social networks do you use?*. Über 90% der Probanden geben an, neben Twitter auch Facebook zu nutzen. Ebenfalls beliebte soziale Netzwerke unter den Teilnehmern sind: Pinterest (37,2%), Google+ (37%) sowie LinkedIn (27,3%). Die Teilnehmer können zudem als Langzeit-Nutzer betrachtet werden (vgl. 5.10 c)), da 74,8% angeben, Twitter seit mindestens einem Jahr zu nutzen. Nur 7,3% der Teilnehmer sind frisch registrierte Nutzer und verwenden die Social-Media-Applikation erst für etwa einen Monat. Abbildung 5.10 d) zeigt die Antworten auf die Frage *Which of the following are common means for you to use Twitter?*. Bemerkenswerterweise geben knapp 90% der Teilnehmer an, Twitter über den Website-Client zu nutzen. Die für das jeweilige Smartphone erhältliche Twitter-App rangiert mit 37,6% nur auf Rang zwei. Dies steht etwas in Diskrepanz zu Twitters offizieller Statistik, dass 82%<sup>17</sup> aller Nutzer-Interaktionen von einem mobilen Endgerät erfolgen. Dabei ist es wohl keine Frage eines Entweder-oder, die die Nutzung Twitters via unterschiedlicher Endgeräte charakterisiert, sondern vielmehr ein Sowohl-als-auch, was bedeutet, dass der Zugang zu Twitter abhängig vom jeweiligen Kontext (Ort, Aufgabe, Tageszeit etc.) sowohl via Website als auch über das mobile Endgerät erfolgen kann. In Bezug auf die angegebenen Twitter-Account-Informationen der Umfrageteilnehmer kann für alle Parameter eine schiefssymmetrische Verteilung mit *long tail* festgestellt werden. Dies lässt auf heterogene Nutzungsweisen des Services schließen. Im Median sendeten die Teilnehmer 118 Tweets ( $Min = 0, Max = 73940, IQR = 1013$ ), folgten 28 Accounts ( $Min = 0, Max = 33000, IQR = 124$ ) und hatten 50 Follower ( $Min = 0, Max = 20000$ ,

<sup>17</sup><https://about.twitter.com/company> Stand: Januar 2017





**Abbildung 5.10.:** Demographische Daten zu a) Bildungsniveau, b) Social-Media-Nutzung, c) Twitter-Nutzungsdauer und d) Zugangsmethode zu Twitter für die gesamte Stichprobe der Twitter-Umfrage

$IQR = 173$ ). Auch bei der Anzahl der favorisierten Tweets verhält es sich ähnlich. Im Median hatte jeder Nutzer 5 favorisierte Tweets ( $Min = 0$ ,  $Max = 55600$ ,  $IQR = 40$ ), wobei aber auch extreme Werte für favorisierte Tweets festgestellt werden konnten. Zwei Probanden gaben an, 55 600 bzw. 8000 Tweets favorisiert zu haben. Diese hohe Anzahl an favorisierten Tweets wird in Abschnitt 5.6.3.3 bei der Motivation [C2] erneut aufgegriffen.

Abschließend lässt sich feststellen, dass sowohl in Bezug auf die demographischen Charakteristika wie Alter, Geschlecht und Ausbildungsniveau, sowie in Bezug auf die Häufigkeit und Art der Twitter-Nutzung eine heterogene und breitgefächerte Stichprobe rekrutiert werden konnte. Ein repräsentatives Sample an Twitter-Nutzern zu rekrutieren, scheint ohnehin nur von Twitter selbst geleistet werden zu können und war nicht als Ziel der Stichprobenzusammenstellung ausgegeben. Vielmehr sollen die subjektiven Eindrücke zu Favorisieren, Aufbewahrens- und Wiederfindenspraktiken von einer Vielzahl an unterschiedlichen Twitter-Nutzern eingeholt werden, was das Sample zu leisten vermag.

## 5.6. Die Rolle der Favoriten-Funktion für das PIM-Verhalten auf Twitter

Um die Frage nach der Motivation und den Gründen für das Favorisieren von Tweets ganzheitlich beantworten zu können, wurden die beiden Freitextfragen:

- *Can you name reasons that lead you to favourite Tweets?*

- *Why did you favourite this Tweet?*<sup>18</sup>

der Twitter-Umfrage qualitativ analysiert. Die erste Frage zielt darauf ab, generelle Erklärungsansätze und Beweggründe für das Favorisieren zu sammeln. Die zweite Frage fokussiert auf den von den Nutzern als letztes favorisierten Tweet und dient dazu, Favorisierungsgründe für ein konkretes Beispiel zu sammeln, um einen vermeintlichen *memory bias* zu überwinden. Dies sollte den Studienteilnehmern dabei helfen, Gründe zu eruieren, die ihnen bei allgemeinen Überlegungen nicht in den Sinn kommen würden. Im Folgenden wird der Auswertungsprozess der beiden Freitextfragen erläutert.

### 5.6.1. Der Auswertungsprozess im Detail

Abbildung 5.11 zeigt den Auswertungsablauf. Die Aufteilung der einzelnen Schritte sowie die Wahl von Begrifflichkeiten ist angelehnt an das Werkzeugkastenmodell nach Margrit Schreier aus Abschnitt 5.4.2.

#### Forschungsfrage

Die Auswertung der Freitextantworten war von Beginn an von einer latenten Forschungsfrage begleitet, nämlich der Frage nach der konkreten Motivation bzw. den Absichten von Twitter-Nutzern, die Favoriten-Funktion zu nutzen. Erste Hinweise, die als Kontextwissen bereits in diesem frühen Stadium der Analyse mit in die zu erwartenden Ergebnisse einfließen, beruhen auf anekdotischer Evidenz aus Blogeinträgen und Foren, in denen bereits erste Favorisierungs-Gründe genannt werden (Mindruta, 2013; O'Neil, 2013). Zudem gibt Twitter in seiner Funktionsbeschreibung mögliche Nutzungskontexte für das Favorisieren von Tweets, wie *like* oder *bookmark*, an (Twitter, 2013). Beide Sachverhalte stehen in Diskrepanz zur Vorgehensweise bei der GT, wie bereits in Abschnitt 5.4.3 beschrieben.

#### Materialauswahl

Für das Erstellen des Kategoriensystems mithilfe der an die GT angelehnten Kodierphasen wurden nicht alle Antworten der Umfrageteilnehmer herangezogen. Aus den insgesamt 331 Antworten wurden 50 zufällig ausgewählt und für den weiteren Kodierprozess genutzt.

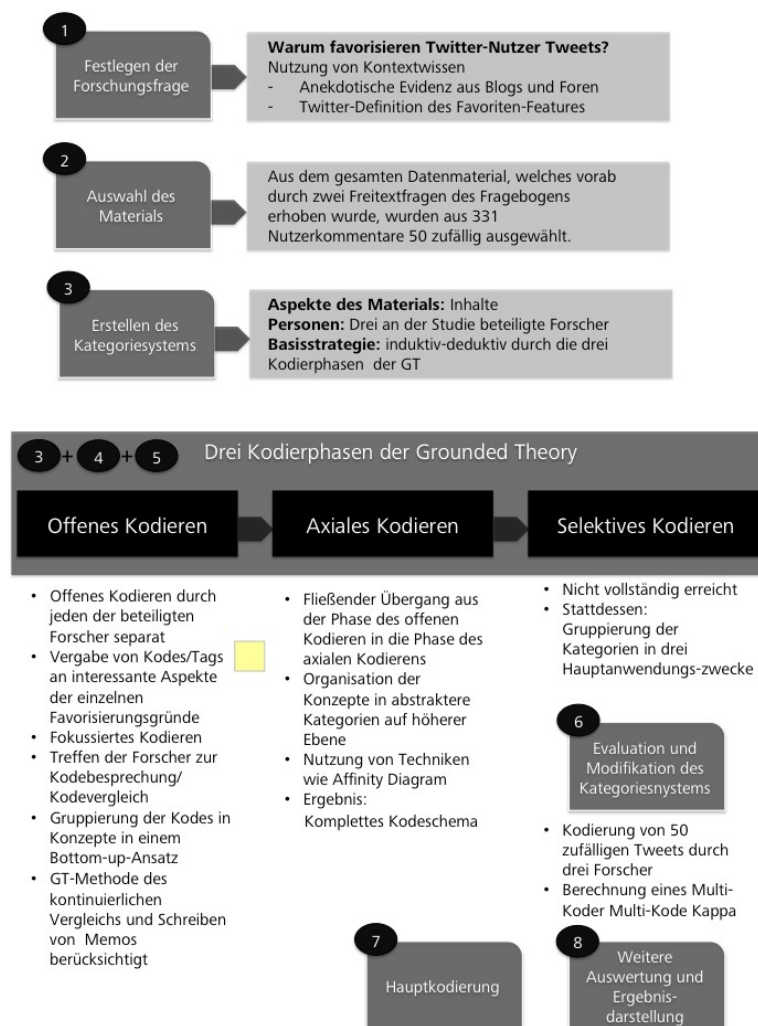
#### Erstellung des Kodierschemas

Orientiert man sich an den Begrifflichkeiten des Werkzeugkastenmodells, so wurden die Materialinhalte (die Gründe) als Kodierobjekte gewählt. Zudem waren drei Personen bei dem Kodierprozess beteiligt. Die Beteiligung von mehreren Personen beim Kodierprozess ist ein Umstand, der in der GT als äußerst wünschenswert erachtet wird und als gute Praxis gilt. Die Basisstrategie des Kodierens kann als induktiv-deduktiv bezeichnet werden, da bereits Vorwissen über mögliche Kategorien besteht und Antworten in diese Kategorien bzw. Codes zugeordnet wurden (deduktiv) aber zusätzlich aus dem Material neue Codes und Kategorien gebildet wurden (induktiv).

#### Kodieren anhand der Phasen der GT

Das ausgewählte Material wurde anhand der Kodierphasen der GT bearbeitet. Am Anfang stand hierbei die Phase des offenen Kodierens. Beim offenen Kodieren werden Tags oder Codes für interessante Aspekte des Materials vergeben. Dies wurde zunächst von allen Beteiligten Personen separat durchgeführt. Anschließend wurden bei einem Treffen alle benutzten Codes zusammengetragen und mit der Methode des ständigen Vergleichs erste Gruppierungen von Codes vorgenommen. In einem Bottom-up-Ansatz entstanden so erste Konzepte. Parallel zu diesem Prozess wurden Memos protokolliert, wie dies in der GT gefordert ist. Abbildung 5.12a zeigt das Ergebnis der ersten Kodierphase. Nach dem offenen Kodieren erfolgte ein fließender Übergang in die Phase des axialen Kodierens, in der versucht wurde, Konzepte in Kategorien höherer Ordnung

<sup>18</sup>Diese Frage war nur in der Umfrage für die Tellwut-Stichprobe enthalten.

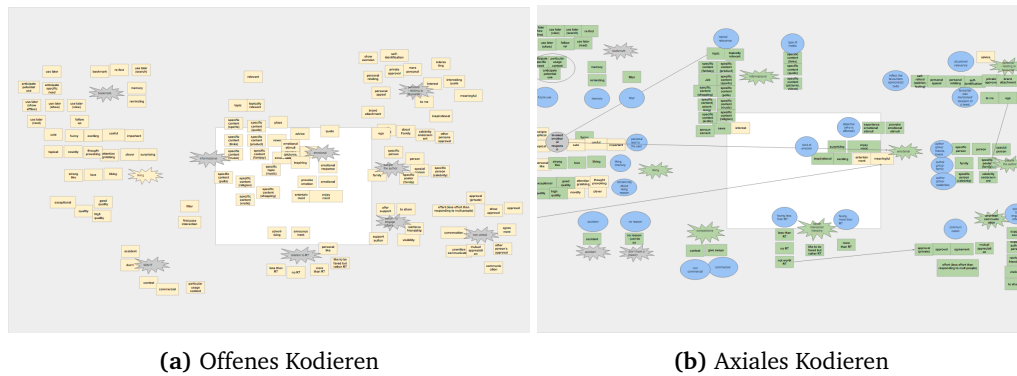


**Abbildung 5.11.:** Visualisierung des Auswertungsprozesses wie er sowohl bei den Favorisierungsgründen als auch bei den Wiederfindensgründen (vgl. Abschnitt 5.8) Anwendung fand. Die einzelnen Schritte entsprechen den Schritten des Werkzeugkastenmodells nach Schreier (vgl. Abbildung 5.9).

zu organisieren. Dies geschah unter der Zuhilfenahme von weiteren Techniken wie etwa der des *affinity diagram*, bei der Kategorien räumlich nach Zusammengehörigkeit angeordnet werden. Zudem wurden die Rohdaten (Codes) stets als Indikatoren für Konzepte gewertet, welche ebenfalls ständig verglichen wurden, um die Bindung zwischen Daten und abstrakteren Konzepten und Kategorien nicht zu verlieren. Abbildung 5.12b zeigt das Ergebnis der Phase des axialen Kodierens. Die räumliche Nähe von Kategorien entspricht ihrer inhaltlichen Ähnlichkeit. Die blauen Ellipsen visualisieren Kategorien höherer Ordnung. Als Ergebnis dieser Phase entstand ein fertiges Kodierschema. In der dritten Phase, der Phase des selektiven Kodierens, wurden die Kategorien weiter in drei Hauptanwendungszwecke eingeteilt.

### Evaluation

Das Kategoriensystem wurde anschließend evaluiert. Hierfür wurden erneut 50 zufällige Antwort-



**Abbildung 5.12.:** Aufteilung der Codes nach den Phasen des Offenen (5.12a) und Axialen (5.12b) Kodierens

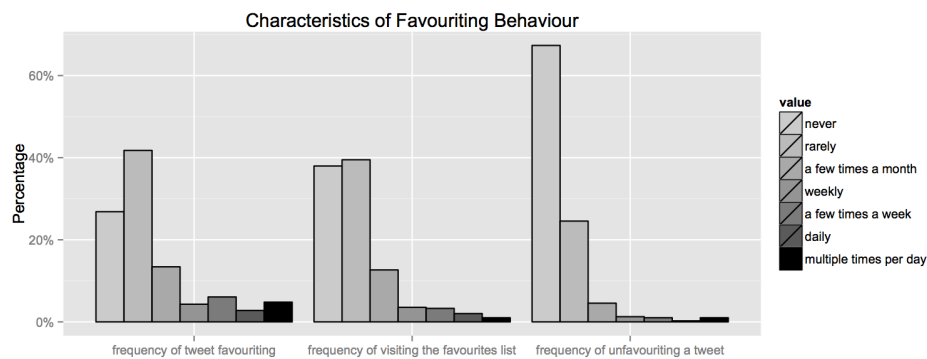
ten aus den 331 Möglichkeiten gezogen und von den drei beteiligten Personen auf der Basis der Codes des Kodierschemas kodiert. Dabei ist anzumerken, dass auf einzelne Antworten mehrere Codes zutreffen können. Um die Übereinstimmung der drei Kodierer auf der Basis von mehreren möglichen Codes zu bestimmen, wurde ein Multi-Kode-Multi-Koder-Kappa berechnet. Hierzu wurde das von Harris and Burke vorgestellte Verfahren zur Bestimmung der Inter-Koder-Reliabilität in R umgesetzt (Harris & Burke, 2005) (vgl. auch 5.4.3). Die Kodierer konnten eine Übereinstimmung von  $\kappa = 0,83$  erreichen, was nach Grouven et al. einer *sehr guten* Übereinstimmung entspricht (Grouven et al., 2007).

Aufgrund der hohen Übereinstimmung wurde das fertige und validierte Kodierschema auf alle 331 Freitextantworten angewendet, wobei das finale Kodieren lediglich vom Autor durchgeführt wurde. Im Folgenden wird das Kodierschema sowie dessen ausführliche Interpretation vorgestellt. Zunächst wird das Favorisierensverhalten jedoch quantitativ beschrieben.

### 5.6.2. Nutzungshäufigkeit der Favoriten-Funktion

Nur 395 Personen (65,1%) der 606 Umfrageteilnehmern geben an, die Favoriten-Funktion zu kennen. Von diesen 395 Personen wiederum antworten 290 Personen (73,5%), die Funktion schon einmal genutzt zu haben. Abbildung 5.13 zeigt die Verteilung der Häufigkeiten der Aktivitäten, die mit Favorisieren in Zusammenhang stehen: die generelle Nutzungshäufigkeit der Funktion, die Häufigkeit des Besuchs der Favoritenliste und die Häufigkeit mit der Tweets entfavorisiert, also aus der Favoritenliste gelöscht werden.

Der Abbildung 5.13 ist zu entnehmen, dass lediglich 31,6% der Teilnehmer angeben, regelmäßig Tweets zu favorisieren. Jedoch muss bemerkt werden, dass zumindest 5% der Probanden täglich bzw. mehrfach am Tag die Favoriten-Funktion nutzen. Die Favoritenliste wird nur selten besucht. 77,4% der Teilnehmer sichten ihre Favoritenliste nie oder nur selten. Ein vergleichbares Bild zeigt das Verhalten der Nutzer beim Entfavorisieren. 91,8% der Nutzer geben an, noch nie einen Tweet aus der Favoritenliste gelöscht zu haben, oder dies nur selten zu tun.



**Abbildung 5.13.:** Häufigkeit des Favorisierens, Entfavorisierens und Favoritenlistenbesuchs im Überblick

### 5.6.3. Motivation und Beweggründe zur Nutzung der Favoriten-Funktion

Tabelle 5.2 zeigt das entwickelte Kodierschema, das 25 unterschiedliche Beweggründe für die Nutzung der Favoriten-Funktion präsentiert.<sup>19</sup> Um die detaillierte Darstellung der einzelnen Codes einfacher zu gestalten, wird jeder Code mit einer eindeutigen Referenznummer versehen. Beispielsweise steht [A2.3] für den Code A 2.3: *Special person as author: celebrity*. Analog hierzu werden auch die von den Teilnehmern gegebenen Gründe mit einer ähnlichen Syntax referenziert, um eine bessere Vergleichbarkeit zwischen unterschiedlichen Gründen und den jeweils angebrachten Codes herstellen zu können. [R1] steht demnach für den ersten Grund. Insgesamt wurden von den Probanden 331 Gründe genannt. Die Zahlen in Tabelle 5.2 geben an, wie häufig ein bestimmter Code verwendet wurde. Ist eine Zahl in Klammern geführt, ist dies der Hinweis wie häufig die übergeordnete Kategorie verwendet wurde, wenn aufgrund fehlender Details in der Beschreibung keine konkrete Zuordnung zu einer Subkategorie bzw. die Verwendung eines Sub-Codes möglich war. Im Durchschnitt wurden pro Grund 1,6 Codes vergeben. Die maximale Anzahl an Codes für einen Grund war sieben. Häufig reichte jedoch die Zuteilung eines einzelnen Codes aus, um die Motivation in der Aussage erfassen zu können. Interessanterweise wurden nicht nur Begründungen für das Favorisieren gegeben sondern auch Erklärungen geliefert, wofür Favorisieren nicht verwendet wird. Begründung [R178] verdeutlicht dies: „I use it like bookmarking, for tweets with interesting links. I do not use it to express approval or enjoyment“.

Als Resultat des selektiven Kodierens können die Favorisierungsabsichten in zwei bzw. drei Hauptnutzungsszenarien eingeteilt werden. Kategorie [A] charakterisiert Favorisieren als Reaktion oder Antwort auf den Inhalt eines Tweets oder dessen Metadaten (Autor). Die offensichtlichste Form einer Reaktion stellt das klassische *Liken* (vgl. Code [A.3]) dar, wobei laut Tabelle 5.2 sehr unterschiedliche Like-Intensitäten seitens der Nutzer vorhanden sein können [A3.1]. Zudem sind die Merkmale eines Tweets, die zum Liken führen, häufig eher subjektiv geprägt (vgl. Code [A3.2]). Aber auch objektive Charakteristika [A3.3], wie eine schöne Satzformulierung, können Like-Gründe sein. Zudem beschränkt sich das Favorisieren nicht nur auf den Inhalt eines Tweets. Auch Metadaten wie der Autor können einen Nutzer zum Favorisieren motivieren. Dies zeigt sich z. B., wenn der Sender eines Tweets ein naher Verwandter ist. Dann kann dies ein Grund für einen Nutzer sein, einen Tweet zu favorisieren (vgl. Code [A2.2]).

Das zweite Nutzungsszenario interpretiert das Favorisieren als Hilfsfunktion oder zielorientiertes

<sup>19</sup>An der Auswertung der Favorisierungsgründe waren Max L. Wilson und David Elweiler mit beteiligt.

<b>A</b>	<b>FAVOURITING AS A RESPONSE/REACTION</b>	<b>384</b>	<b>B</b>	<b>FAVOURITING FOR A PURPOSE/AS FUNCTION</b>	<b>144</b>
A1	Tweet is informational	(3) 58	B1	Bookmarking	(4) 75
A1.1	Topically relevant	32	B1.1	Future use	52
A1.2	Relevant (Multi-) Media	23	B1.1.1	Future interaction step	47
			B1.1.2	Anticipated need	5
A2	Special people as author	(12) 31	B1.2	Memory	19
A2.1	Friends/lovers	5	B1.2.1	Encoding step	1
A2.2	Family	3	B1.2.2	Memento	18
A2.3	Celebrity	11			
			B2	Unwritten Communication	40
A3	Liking	(2) 206	B2.1	Agreement or approval (private)	27
A3.1	Liking intensity	19	B2.2	Engaging in conversation	7
A3.2	Subjective liking reason	172	B2.3	Trying to engage others	6
A3.3	Objective liking reason	13			
			B3	Competitions	(5) 13
A4	Personal relating	49	B3.1	Non-commercial	2
A4.1	Situationally relevant	7	B3.2	Commercial	6
A4.2	Tweet reflects the authors opinion	28			
A4.3	Favouriter was mentioned/recipient	14	B4	Twitter functions relationship	(1) 16
			B4.1	Favouriting means less than RT	13
A5	Emotional stimulus	40	B4.2	Favouriting means more than RT	2
A5.1	Objective	15			
A5.2	Kind of emotion	25	C	NO REASON BUT INTERESTING BEHAVIOUR	9
			C1	Accident	3
			C2	No specific reason	6

**Tabelle 5.2.:** Kategoriensystem/Kodierschema der Favorisierungsgründe als Resultat der qualitativen Datenauswertung

Handeln, wobei Favorisieren einen bestimmten Zweck erfüllt oder dazu genutzt werden kann, um ein Ziel zu erreichen. Die Nutzung der Favoriten-Funktion, um ein Lesezeichen für einen bestimmten Tweet zu setzen (vgl. Kode [B.1]) und diesen einfach aufzubewahren und leichter zugänglich zu machen, ist ein Beispiel für ein solches Ziel. Innerhalb dieser Kodedimension sind die tatsächlichen Beweggründe noch feiner ausdifferenzierbar. Eine weitere Funktion ist Favorisieren als non-verbale Kommunikation [B2]. Auch hier erfüllt das Setzen eines Sternchens einen bestimmten Zweck: die Vermittlung einer Botschaft. Schließlich lässt sich noch eine dritte Gruppe an Motiven für das Favorisieren ausmachen. Diese beschreiben weniger tatsächliche Gründe, als vielmehr besonderes Verhalten und sind unter Kategorie [C] geführt. Im Folgenden werden die einzelnen Beweggründe und Motive anhand der Codes und ausgewählter Beispielgründe erläutert.

### 5.6.3.1. Kodekategorie [A]: *Favouriting as a response/reaction*

#### **Tweet is informational [A1]**

Spezielle Interaktion mit einem Tweet (Weiterleiten und/oder Favorisieren) wird in der Literatur als Zeichen für die Relevanz eines Tweets gewertet (vgl. Abschnitt 3.2.1.1). Der informationelle Mehrwert (vgl. Kode [A1]) bzw. die inhaltliche Relevanz eines Tweets (vgl. Kode [A1.1]) sind Motive, die auch in dieser Umfrage als Favorisierungsgründe angegeben wurden. Eine weitere Differenzierung kann durchgeführt werden, wenn Nutzer sich explizit dazu geäußert haben, dass es das Video, das Bild oder der Link waren, die einem Tweet angehängt sind und zur Favorisierung führen (vgl. Kode[A1.2]). Inhaltliche Relevanz lässt sich in Nutzerkommentaren anhand der in ihnen erwähnten spezifischen Themen wie Sport („CRICKET MATCHES“ [R206]), Shopping („People put on good shopping sites“ [R133]), Musik („I was excited for nsync to come back!“ [R327]) oder Politik („Interesting news, Tweets for political figures that I agree with, Statements that I agree with“ [R193]) festmachen. Die Relevanz des jeweiligen Anhangs lässt sich aus folgenden Nutzerkommentaren ableiten: „They have useful links in them“ [R119], „Interesting articles, blogs, pictures, videos, memes (obviously)“ oder „It list url that in line with my interest“ [R102]. Obwohl Ähnlichkeiten bestehen, können diese Gründe klar von Gründen getrennt werden, die Favorisieren als Bookmarking interpretieren. Diese Begründungen enthalten zusätzlich Schlüsselphrasen wie *save*

for later, remember oder go back und nennen zudem das zukünftige Vorhaben (Kodes [B1.1] *Future use*) mit dem aufbewahrten Tweet. Favorisieren zum Setzen eines Lesezeichens (Kode [B1]) wird in Abschnitt 5.6.3.2 erläutert.

### Special people as author [A2]

Treffen Nutzer beim Lesen ihrer *Timeline* auf einen Tweet von einem bestimmten Autor, ist dies ein Beweggrund diesen Tweet zu favorisieren. Eine Rolle spielt dabei auch die Autor-Rezipient-Beziehung. Innerhalb der Nutzerkommentare können drei Gruppen identifiziert werden, auf deren Tweets Nutzer mit Favorisieren reagieren. [A2.1] *Friends and Lovers* („Something that a friend posts that i really like or agree with.“ [R24]), [A2.2] *Family* („Because my god daughter tweeted it“ [R288]), [A2.3] *Celebrity* („Shoutout from a favorite celeb“ [R242]). Kode [A2.3] steht dabei in sehr engem Zusammenhang mit Kode [A4.3] *Favouriter was mentioned/recipient*, da Nutzer dann favorisieren, wenn sie eine Antwort (@Reply) oder Erwähnung (@Mention) von einer Berühmtheit erhalten. Erkennbar ist dies an Kommentaren wie „Because it was a reply from robert englund and freddy krueger is my favorite horror movie villain of all time.“ [R295] oder „[...] comment to me by a celeb“ [R74]. Auffällig ist die hohe Anzahl an Kommentaren (12), bei denen offensichtlich ist, dass der Grund für das Favorisieren eine spezielle Autor-Nutzer-Beziehung ist, jedoch aus dem Kontext des Kommentars keine explizite Beziehung abgeleitet werden konnte. Kommentare, die mit einem unspezifischen Kode [A2] versehen wurden, sind z. B. „It was directly to me by one of my favorite people“ [R227]. Es ist offensichtlich, dass sowohl die direkte Erwähnung (Kode [A4.3]) als auch die Rolle der Person (Kode [A2]) das Favorisieren motivieren. Die besondere Rolle eines Autors/Senders wird auch bei Counts und Fisher evident, die feststellten, dass ein deutlich besseres Erinnerungsvermögen an Tweets besteht, die von bestimmten Personengruppen gesendet wurden (Counts & Fisher, 2011). Ein Fakt, der immer wieder in der Arbeit angesprochen werden wird, da er auch für das Wiederfinden von Bedeutung ist.

### Liking [A3]

Das am häufigsten genannte Motiv für die Nutzung der Favoriten-Funktion stellt der *like* dar. 206 der 331 Freitextantworten enthielten *liken* als Motivation für Favorisieren. Dies ist nicht verwunderlich, da dieses Konzept auch in vielen anderen Social-Media-Applikationen verbreitet ist und Nutzer das Favorisieren mit der Like-Funktion von Facebook assoziieren. Dies zeigen Kommentare wie „Generally use it analogously to 'like' on facebook“ [R182] oder „[...] almost like liking it on Facebook“ [R190]. Zudem wird der Like von Twitter als ein möglicher Hauptanwendungszweck der Funktion kommuniziert (Stone, 2006).

Betrachtet man die Nutzerantworten im Detail, kann ein Like weiter ausdifferenziert werden. Hierbei wird ersichtlich, dass Tweets favorisiert werden, wenn eine bestimmte Intensität [A3.1] des Mögens seitens des Nutzers erreicht wird. Favorisieren besitzt somit in Abhängigkeit vom Nutzer einen niedrigeren oder höheren Grad an Zuneigung, der, wenn er überschritten wird, zu einem Signal via Favorit führt. Diese Intensitätssteigerung veranschaulichen diese drei Kommentare: „If i really like them, i will favourite“ [R95], „When I like them a lot“ [R204] und „I love mcdonalds“ [R230]. Neben der Intensität des Mögens werden auch explizit Charakteristika oder Merkmale genannt, die zum Like führen. Diese können subjektiver [A3.2] oder objektiver [A3.3] Natur sein. Aussagen wie „It was well said[...]“ [R304] oder „Thought it was funny and a pun“ [R308] deuten auf eine gewisse sprachliche Versiertheit hin, die objektiv feststellbar ist. Die Begründung „Funny“, die durchaus sehr häufig auftritt, kann hingegen lediglich als subjektive Begründung [A3.2] gesehen werden, da verschiedene Personen unterschiedliche Schwellen dafür haben, was sie als lustig empfinden. Vergleichbare Gründe für das Favorisieren sind „I find the tweet funny or the tweet announces something that I like“ [R76] oder „When they are sweet“ [R56]. Auch hier werden subjektive Begründungen gegeben, die die Verwendung des Kodes [A3.2] erfordern. Hierbei muss differenziert werden zwischen dem Tweetinhalt und dem emotionalen Zustand, den ein Tweet bei

einem Nutzer auslöst (vgl. Kategorie [A5]: *Emotional Stimulus*). Häufig wurden als Begründungen für ein Like auch die Metadaten eines Tweets, wie der Autor, herangezogen. Kode [A2]: *Special person as author* trägt dieser Begründung Rechnung und wurde bereits weiter oben erläutert.

#### **Personal Relating [A4]**

Die Kategorie [A4] vereinheitlicht Favorisierungsgründe unter dem Konzept des persönlichen Bezuges zwischen Tweet und Favorisierenden. Kode [A4.1] beschreibt dabei Favorisieren aufgrund der situativen Relevanz eines Tweets, wie sie beispielsweise ein guter Ratschlag besitzen kann. Situative Relevanz wird von Cosijn und Ingwersen wie folgt definiert: „the relation between the perceived situation, task or problem at hand and the usefulness of information objects as perceived by the user“ (Cosijn & Ingwersen, 2000, S.539). Informationsobjekte, die von situativer Relevanz sind, helfen folglich einer Person bei der Entscheidungsfindung oder der Problemlösung in einer bestimmten Situation. Kommentare die Kode [A4.1] erhalten, beinhalten Stichwörter wie *important* oder *relevant* mit einem bestimmten Zeitbezug. Kode [A4.2] beschreibt Favorisieren als Reaktion auf einen Tweet dessen Inhalt, Einstellung oder Meinung mit der des Favorisierenden übereinstimmt. Beispielkommentare, die dies verdeutlichen sind: „messages that supported what I believed and felt at the time“ [R220], „Sounds like something I would say about the topic“ [R265] oder „I did it because it was an opinion that i agreed with“ [R233]. Es wird deutlich, dass Favorisieren hier sowohl als Reaktion als auch als Funktion interpretiert werden kann, weshalb Begründungen, die mit Kode [A4.2] markiert werden können, zusätzlich den Kode [B2.1] *Agreement or approval (privacy)* erhalten. Kode [B2.1] wertet Favorisieren als Form non-verbaler bzw. nicht-textueller Kommunikation. Zusätzlich kann der Nutzer eine gewisse Art an Selbstidentifikation mit dem Tweetinhalt erfahren. Inhalte solcher Art werden vor allem auch durch Unternehmen oder Marken via Twitter kommuniziert (vgl. [R65] „a brand that fits my way of life“). Erreicht der persönliche Bezug ein ausreichend hohes Level, kann durchaus ein emotionaler Zustand beim Nutzer ausgelöst werden, welcher dann die Verwendung eines Kodes der Kategorie [A5] *Emotional stimulus* verlangt. Kode [A4.3] *Favourite was mentioned/recipient* stellt den am stärksten Ich-bezogenen Kode aller Kategorien dar. Tweets werden dann von Nutzern favorisiert, wenn diese an sie gerichtet sind oder sie in diesen Tweets erwähnt werden. Diese narzisstische Haltung wird vor allem in Begründungen evident wie: „Because they mentioned me“ [R237] oder „A very interesting reply to me“ [R5]. Ein ego-zentrierter Favorit wurde auch bereits bei Kategorie [A2] vorgestellt. Dort gewinnt die Antwort auf einen Tweet zusätzlich an Bedeutung, wenn sie von einer beliebten Berühmtheit stammt.

#### **Emotional Stimulus [A5]**

Aus Konzept [A5] geht hervor, dass Tweets favorisiert werden, wenn sie den emotionalen Zustand des Lesers beeinflussen. Dies kann einerseits auf einem globalen Level für eine Vielzahl von Nutzern der Fall sein, oder auf einem individuellen Level für einzelne Personen gelten, wenn der persönliche Kontext eine Rolle spielt. Favorisierte Tweets wurden als „inspiring“ [R145], oder „Made me feel special“ [R245] beschrieben. Dass Tweets eine Stimmungsänderung herbeiführen können und dadurch favorisiert werden, wird auch in folgender Begründung ersichtlich: „They make me love or have some kind of emotional meaning resonating with me.“ [R200]. Da die Art der Emotion, die beim Nutzer hervorgerufen werden kann, sehr unterschiedlich ist, bildete sich während der Datenauswertung ein eigener Kode [A5.2] heraus. Umgekehrt werden Tweets favorisiert, um den Gefühlszustand des Autors zu beeinflussen. Dies belegen Kommentare wie „Wanted to make her laugh“ [R255] oder „It’s an admirable milestone, i wanted to grand her recognition so she would feel like people recognize her accomplishment“ [R287]. Die Kodes der Kategorie [A5] treten logischerweise nicht alleine auf, sondern werden je nach Detailgrad in den Kommentaren von inhaltlichen Kodes wie [A1.1] und [A1.2] begleitet.



### 5.6.3.2. Kodekategorie [B]: *Favouriting for a purpose/as a function*

#### Bookmarking [B1]

Die Nutzung der Favoriten-Funktion als Lesezeichen-Funktion stellt das zweite offizielle Nutzungsszenario der Funktion dar. Sie ist somit die einzige Funktion — zumindest zum Zeitpunkt der Studie — im Twitter-Universum, die das Aufbewahren und das Management von Tweets erlaubt.

Für die Fragestellung nach dem Wiederfindensverhalten ist *Bookmarking* [B1] das wichtigste Konzept, das sich aus dem Kodierprozess herausbildet. Kodes dieser Kategorie fanden am zweithäufigsten Verwendung — insgesamt 75 Mal (22%) — was deutlich macht, dass Nutzer das Bedürfnis besitzen, Tweets wiederzufinden und hierfür auch Schritte unternehmen, die Wiederfinden potentiell erleichtern, indem durch das Favorisieren Tweets in einem separaten Speicher aufbewahrt werden. Nutzer geben explizit an, Favorisieren als Bookmarking-Funktion zu nutzen: „Bookmarking articles or videos to review later when I have more time Useful for my professional activity“ [R194], „I use the favorite feature like a bookmark feature; tweets i like so i can see them again later.“ [R100].

Bookmarking per se kann jedoch sehr unterschiedlich motiviert sein: zum einen durch einen zukünftigen Interaktionsschritt (vgl. Kode [B1.1] *Future use*) zum anderen als Gedächtnisstütze (vgl. Kode [B1.2] *Memory*). Beides sind unterschiedliche Motive, welche durch die entsprechenden Kodes [B1.1] und [B1.2] repräsentiert sind. Kommentare, die auf zukünftige Interaktionsschritte hindeuten, enthalten häufig Schlüsselwörter wie *again*, *later* oder *go back*. Neben diesen Schlüsselwörtern werden häufig auch konkrete Aktivitäten genannt, die beim Schritt des Favorisierens eine zukünftige Handlung wie *to read*, *to show*, *to view*, *to search* oder explizit *to re-find*, beinhalten. Beispielkommentare veranschaulichen dies: „I want to read them again“ [R161], „To save it for later (for example a link to something I want to view later)[...]“ [R190] oder „[...] I generally favourite things that I think I will want to re-find again in the future. This is extremely difficult to do using the Twitter search system“ [R180]. Das Motiv [R180] zeigt nicht nur, dass Wiederfinden ein auftretendes Nutzerbedürfnis darstellt, sondern weist auch darauf hin, dass Wiederfinden mithilfe der Twitter-Suche kaum möglich ist bzw. Wiederfindensaktionen von dieser nur schlecht unterstützt werden. Auch wenn einige Kommentare eher vage Aussagen enthalten, wofür Tweets später genutzt werden, gibt es auch Motive, die konkret begründen — durch ein antizipiertes Informationsbedürfnis — wofür ein Tweet in der Zukunft eingesetzt wird: „I am a foodie and a fitness fanatic! i favorite workout routines or recipes to try later. also, motivational pics and sayings to inspire me when i need it“ [R62] oder „So i can use it when i work out“ [R228].

Die zweite Kategorie [B1.2] *Memory* deckt ebenfalls weitere Motive ab. Zum einen geben Nutzer an, Tweets zu favorisieren, um sich besser an den Tweet erinnern zu können. Favorisieren käme somit einem zusätzlichen Schritt beim Einprägen des Tweets gleich (vgl. Kode [1.2.1]: „I use favourites as bookmarks, to remind me to look up a tweet or a topic in more detail later.“ [R179]. Ähnliche Effekte der Erinnerungshilfe durch zusätzliche Interaktionsschritte können Elsweiler, Bailie und Ruthven beim Ablegen von E-Mails in Ordnern beobachten (Elsweiler et al., 2008). In Nutzerkommentaren treten häufiger Schlüsselwörter wie *remembering* oder *reminding* auf, jedoch kann nicht in jedem Fall eine Absicht zur Einprägung des Tweets unterstellt werden. Zum anderen werden Tweets favorisiert, um eine Liste beliebter Tweets zu erstellen und das Andenken an diese Tweets zu bewahren. Motive wie „So i can always remember the tweet“ [R13] oder „Quotes or links that i want to remember“ [R133] geben Hinweise darauf, dass Nutzer eine Sammlung interessanter Tweets anlegen, die sie häufiger besuchen als nur für eine weitere Interaktion, wie dies bei Kode [B1.1.1] der Fall ist.

Es wird deutlich, dass die Favoritenliste unterschiedliche Nutzerbedürfnisse zu erfüllen hat. Einerseits die Rolle des Kurzzeitspeichers für Nutzer, die einen Tweet wiederfinden wollen, und anderer-

seits eine Liste zur Langzeitarchivierung zum Aufbau von Mementos der schönsten Tweets. Erstes würde erklären, weshalb bestimmte Nutzer, obwohl sie angeben, häufig Tweets zu favorisieren, keine Tweets in ihrer Favoritenliste besitzen, da nach dem Wiederfinden der Tweet entfavorisiert wird oder Tweets in regelmäßigen Abständen bündelweise gelöscht werden. Ein Verhalten, das aus dem PIM-Verhalten mit E-Mails bekannt ist und von Whittaker und Sidner als *spring cleaning* (Whittaker & Sidner, 1996) bezeichnet wird und bereits in Abschnitt 2.2.3 beschrieben wurde.

### Unwritten Communication [B2]

Kode [B2] umfasst das Konzept der Zustimmung oder Unterstützung durch Favorisieren als nicht-textuelle Form der Kommunikation, was durch das Senden eines Tweets gemacht werden könnte. Dies wird auch explizit in Nutzerkommentaren deutlich: „[...]sometimes just as an approval of an answer without having to write 'i agree'“ [R168]. Vergleichbar ist dies mit einem Nicken in einer Face-to-Face Unterhaltung als Form non-verbaler Kommunikation (vgl. Kode[B2.1]). Nutzer begründen die Nutzung der Favoriten-Funktion für diesen Zweck mit Kommentaren wie „I did it because it was an opinion that i agreed with“ [R234] oder „Because i agree with the author“ [R15]. Auf eine schriftliche Antwort wird verzichtet bzw. wird diese als entbehrlich charakterisiert: „I asked a question, more than one person answered, and i didn't want to answer either of them so i just favoured their tweets.“ [R256]. Auch das Konzept von Favorisieren als Benachrichtigungsfunktion, welches unter Kode [B2.2] erläutert wird, ist hier zu erkennen. Obwohl eine vermeintliche Ähnlichkeit zwischen Kode [A3] *Liking* und [B2.1] *Agreement or approval (privacy)* besteht, ist ein Unterschied zwischen den Kodes zu erkennen, der sich auf die Frage nach dem Adressaten oder Empfänger des Likes bezieht. Letztendlich ist bei einem Like stets der Ausführende selbst der Adressat, bei Favorisieren aufgrund der Zustimmung ist das Favorisieren als Signal an den Autor bzw. Sender des Tweets gerichtet. Über einfache Zustimmung hinaus interpretieren Nutzer das Favorisieren als eine privatere Form der Zustimmung im Gegensatz zu Retweeten, welches von einem Großteil der Nutzer als öffentliche Zustimmung interpretiert wird. Dies wird in folgenden Kommentaren deutlich: „[...]3.To signal approval of a tweet that I don't want to retweet to my followers“ [R175] oder „[...]to show support of a tweet that doesn't really apply to me (so I don't want to retweet it[...])“ [R190]. Die private Form der Zustimmung ist bei Nutzern ein wichtiger Punkt, da sie vermeiden wollen, Tweets an ihre Follower weiterzusenden, wenn sie glauben, dass diese Tweets für sie nicht angebracht oder uninteressant sind. Kode [B2.1] steht dabei in sehr engem Zusammenhang mit Kodes bzw. dem Konzept aus [B4] *Twitter-functions relationship* und der Frage, wie unterschiedliche Nutzer die Bedeutung von Twitter-Funktionen interpretieren und deren Beziehung zueinander wahrnehmen. Neben non-verbaler Kommunikation wird das Favorisieren von Nutzern auch als Mittel genutzt, um Kenntnisnahme und Aufmerksamkeit gegenüber dem Sender zu kommunizieren und somit als eine Art Benachrichtigungsfunktion verwendet, um zu zeigen, dass ein bestimmter Tweet gelesen wurde. Nutzer kommunizieren durch Favorisieren ihre Präsenz als Zuhörer. Folgende Nutzerkommentare belegen dies: „show people that i know that I've read and appreciated their tweet[...]“ [R171], „[...]To signal that I saw a particular tweet from a person that I know well“ [R175] oder „I wanted the tweeter to know i read and liked it“ [R232].

Die letzte Dimension des Konzeptes [B2] interpretiert Favorisieren als eine Möglichkeit, andere Nutzer zu beteiligen und nicht nur selbst Beteiligung zu kommunizieren. Kode [B2.3] erwächst aus Gründen wie „support an event“ [R20], „to get people to join“ [R271], „to spread the word“ [R331] und „i wanted to share it with someone“ [R246]. Bei den letzten beiden wird offensichtlich, dass Fehlinterpretationen über die Funktionalität bzw. das Ergebnis von Favorisieren vorliegen, da favorisierte Tweets nicht automatisch für andere sichtbar sind, außer der Nutzer erwartet, dass andere Nutzer dessen Favoritenliste betrachten.

### Competitions [B3]

Kategorie [B3] repräsentiert Nutzerkommentare, die als Begründung für das Favorisieren die Teil-

nahme an Wettbewerben oder Preisausschreiben angeben. Diese Kategorie ist sehr eindeutig und bietet wenige Überschneidungen mit anderen Kategorien. Bei Wettbewerben kann unterschieden werden zwischen kommerziellen Wettbewerben, also das Gewinnen von realen Sachpreisen wie Tickets zu einem Sportereignis, oder nicht-kommerziellen Wettbewerben auf Plattformen wie Favstar.de<sup>20</sup>, auf denen Nutzer das Ziel verfolgen, Tweets zu verfassen, die von den Followern und der Community möglichst viele Favorisierungen erhalten. Eine Differenzierung zwischen kommerziellen und nicht-kommerziellen Motiven ist jedoch oft nicht durchführbar, da der tatsächliche Kontext in den Kommentaren fehlt: „to try to win a contest“ [R244] oder einfach „for points“ [R273]. Bei einigen Begründungen deuten jedoch Schlüsselwörter wie *giveaways* oder *freebies* auf einen kommerziellen Kontext hin. Beispiel für Kategorie [B3.1] und damit einem kommerziellen Kontext sind: „I did because i love giveaways“ [R316] und „Who wouldn’t want to win world series tickets?“ [R249].

#### **Tweet interaction hierarchy [B4]**

Das Konzept [B4] behandelt weniger Favorisieren als Funktion, sondern fokussiert auf das Verhältnis zwischen Favorisieren und anderen Twitter-Funktionen, im Besonderen dem Weiterleiten (RT) von Tweets und welche Gründe Nutzer besitzen, zwischen einer Funktion zu wählen, oder sich dafür entscheiden, beide zu verwenden. Wie bereits bei Kode [B2.2] deutlich wurde, interpretieren Nutzer das Favorisieren als eine privatere Form, um Zustimmung zu einem Tweet zu kommunizieren und entscheiden sich somit pro Favorisieren und gegen das Weiterleiten eines Tweets, was in den Augen der Nutzer einer öffentlichen Zustimmung gleich kommen würde: „Most often if they’re funny/amusing/interesting. Usually in a way that makes me laugh but I don’t retweet them cause I don’t think others might find them funny (often because they’re offensive / too geeky / too niche).“ [R183]. Nutzer scheinen ebenfalls ein gewisses Wertesystem für Tweets zu entwickeln, das ihnen bei der Entscheidung über Folgeaktionen mit einem Tweet hilft. Dieses Wertesystem führt eine Interaktionshierarchie ein, da laut Nutzerkommentaren bestimmte Tweets nicht den Wert besitzen, um weitergeleitet (RT) zu werden. Folgende Begründungen belegen diese Hierarchie: „Funny/interesting but not worthy of a RT“ [R205], „I feel they’re perfectly expressed but would not make sense if i retweeted them, which is what i usually do instead“ [R136] oder „It was well said but not worth retweeting. it was too late to retweet“ [R304]. Die Begründung [R304] lässt darüber hinaus auch auf eine zeitliche Dimension schließen, die bei der Wahl zwischen den Funktionen eine Rolle spielt. Umgekehrt gibt es auch Nutzer, die dem Favorisieren eine höhere Bedeutung und somit mehr Wert zusprechen (vgl. Kode [B4.2]: „If it’s a very good tweet, that’s wort a RT or more[...]“ [R191] und „I like them more than a retweet“ [R122]. Diese Wertschätzung von Favorisieren tritt mit zwei vergebenen Kodes aber relativ selten auf. Schließlich existieren neben Hinweisen für eine hierarchische Beziehung zwischen Favorisieren und Weiterleiten auch Hinweise, die auf eine sequentielle Abhängigkeit zwischen den Features schließen lässt. Eine denkbare Interaktionskette könnte aus einem vorherigen Favorisieren mit späterem Wiederfinden und einem anschließenden Weiterleiten bestehen. Favorisieren fungiert somit als eine Art Vorverarbeitungsschritt für ein späteres Weiterleiten. In [B4] wird evident, dass Favorisieren mit mehr Privatheit und einer geringeren Wertattribution charakterisiert ist.

#### **5.6.3.3. Kodekategorie [C]: *No reason but interesting behaviour***

Trotz einer Vielzahl an heterogenen Absichten für das Favorisieren finden sich unter den Fragebogenantworten auch Kommentare, die keine Beweggründe für das Favorisieren äußern, jedoch unerwartetes und somit interessantes Verhalten beschreiben. Hierzu zählt zum einen das unabsichtliche Favorisieren [C1] und zum anderen das Favorisieren ohne speziellen Grund [C2].

<sup>20</sup><http://de.favstar.fm/>

**Accident [C1]**

Kommentare wie „Usually accidental“ [R96], „no special. by accident from todays perspective (-;“ [R163] oder „Accidental! but nice as tellwut noticed“ sind Kommentare, welche unabsichtliches Favorisieren beschreiben und mit [C1] *Accident* gekennzeichnet wurden. Sie weisen auf ein mögliches Usability-Problem mit dem Twitter-User Interface hin. Aufgrund der Tatsache, dass die Icons für Reply, RT, FAV und andere Funktionen sehr nahe beisammen liegen und auch sehr klein sind, werden diese häufig auch unabsichtlich geklickt. Ob die Nutzer hier eine andere Funktion verwenden wollten, kann nur spekuliert werden. Vermutlich ist diese Form des unbeabsichtigten Favorisierens auch gehäuft im mobilen Kontext der Fall, da dort zum Downscroll der *Timeline* immer auch ein Touch-Event auf Tweets selber erfolgen muss.

**No specific reason [C2]**

Kommentare, die mit Kode [C2] markiert wurden, sind Motive, die auf den ersten Blick als schlechte Fragebogenantworten abgetan werden könnten. Da diese jedoch auch häufiger auftraten, konnte ein Muster erkannt werden, das Favorisieren ohne tatsächlichen Grund evident werden lässt. Nutzer geben an, keinen konkreten Grund zu haben („I don't have a reason“ [R60]) bzw. nennen zu können oder einfach so („just because“ [R251]) Tweets zu favorisieren. Das Favorisieren kann hier als eine Art Übersprungshandlung interpretiert werden oder als Handlung, die einen Tweet als gelesen markiert, wie dies auch bei E-Mails der Fall ist. Letztendlich kann das Favorisieren ohne einen konkreten Grund die Begründung für die hohe Anzahl (55 600 und 8000) an favorisierten Tweets mancher Nutzer sein, wie in Abschnitt 5.5 angeführt. Bei einer so hohen Anzahl an favorisierten Tweets kann man davon ausgehen, dass vermeintlich alle Tweets der *Timeline* eines Nutzers favorisiert werden, ohne jeweils spezielle Gründe für jede einzelne Favoriten-Aktion zu haben.

**5.6.3.4. Diskussion der Favorisierungsgründe**

Das Kodierschema zeigt, dass Nutzer eine Vielzahl an heterogenen Gründen für die Nutzung der Favoriten-Funktion besitzen. Mehrere Konzepte von Favorisieren koexistieren parallel, was sich in Nutzerkommentaren widerspiegelt, in denen häufig mehrere Begründungen gleichzeitig enthalten sind. Einen Extremfall stellt dabei der Kommentar [R48] dar, welcher mit sieben Codes ausgezeichnet wurde: „To end a conversation, or if something is funny but i don't want to/can't (p/a) rt it. also if someone's having a shitty day and tweet about it, you can like it to know you're sorry abt that“. Zunächst nennt der Kommentar die Favoriten-Funktion als Mittel zum Beenden einer Konversation, also als Form non-verbaler Kommunikation (Kode [B2.1]). Gleichzeitig wird durch die Passage „you can like it to know you're sorry“ Aufmerksamkeit gegenüber dem Sender signalisiert (Kodes [B2.3] und [B2.2]), wobei auch eine emotionale Komponente eine Rolle spielt, was die Verwendung von Kode [A5.2] erfordert. Der emotionale Stimulus soll aber auch bewusst beim Rezipienten des Favoriten erzeugt werden, was letztlich in der Verwendung von Kode [A5.2] mündet. Zusätzlich wird Favorisieren als *like* interpretiert, wobei explizit lustige Tweets als Begründung angegeben werden (Kode [A3.2]). Zudem wird eine Aussage über das Verhältnis zwischen Favorisieren und Retweeten getroffen [B4.1]. Favorisieren basiert folglich auf Nutzerbedürfnissen, deren — wenn sie überhaupt durch das UI unterstützt werden — Unterstützung nur durch ein Icon erfolgt, also schlecht oder gar nicht unterstützt werden. Die Tatsache, dass in einzelnen Nutzeraussagen derart viele Gründe enthalten sind, spricht gegen die Annahme von Gorrell und Bontcheva, dass Nutzer, wenn sie Tweets favorisieren, dies stets aus einem Grund tun (vgl. Abschnitt 3.2.2.2). Insgesamt kann festgestellt werden, dass die Frage bezogen auf den zuletzt favorisierten Tweet wesentlich konkretere Einblicke in die Gründe der Nutzer bot als die Frage nach den generellen Favorisierungsgründen.

Nicht nur durch den Vergleich der Kommentare verschiedener Nutzer, sondern auch innerhalb der

Aussagen einzelner Nutzer wird deutlich, dass Favorisieren für eine Bandbreite an unterschiedlichen Funktionen und non-verbalen bzw. non-textuellen Kommunikationsmechanismen verwendet wird. Eine wesentliche Erkenntnis des Kodierschemas ist die Tatsache, dass viele Gründe und Motive zur Nutzung der Retweet-Funktion die von boyd, Golder und Lotan identifiziert wurden, sich auch bei der Motivation für das Favorisieren erkennen lassen (boyd et al., 2010). Beispielsweise werden beide Features dazu genutzt, um Aufmerksamkeit in einer Unterhaltung zu signalisieren, Freundschaften zu verstärken, Berühmtheiten zu huldigen oder narzisstische Züge auszuleben. Was bei boyd, Golder und Lotan als *Ego-RT* charakterisiert ist, wird beim Favorisieren durch Kode [A4.3] deutlich. Es ist interessant zu erkennen, dass, obwohl viele ähnliche Motive für die Verwendung beider Funktionen bestehen, die Nutzer den Funktionen unterschiedliche Aussagekraft zuweisen. Nutzer besitzen offensichtlich unterschiedliche mentale Modelle der beiden Funktionen Favorisieren und Weiterleiten. Retweeting führt dazu, dass Tweets erneut an die eigenen Follower gesendet werden. Obwohl dies nicht unbedingt ein Zeichen von Zustimmung sein muss, bedeutet das Weiterleiten, dass man die Information selber für angemessen genug erachtet. Favorisieren hingegen — und das wurde bereits bei Kode [B4] angesprochen — wird als eine private Form der Zustimmung interpretiert, da diese Aktion nicht gleich zu einer massenhaften Verbreitung der jeweiligen Information führt, sondern vermeintlich nur für einen selbst sichtbar ist. Hiergegen spricht wohl die öffentlich zugängliche Favoritenliste bei der auch die als privat vermuteten Zustimmungen noch sichtbar sind. Favorisieren als privatere Form der Zustimmung und die öffentliche Zugänglichkeit der Favoritenliste stehen sich folglich konträr gegenüber. Interessant wäre zu analysieren, inwiefern sich weitergeleitete und favorisierte Tweets in dieser Hinsicht unterscheiden und ob durch die Favoritenliste latente Neigungen und Meinungen eines Nutzers offenkundig gemacht werden können. Dass Favoritenlisten als Form der Selbstdarstellung dienen können, wird von Marwick und boyd diskutiert (Marwick & boyd, 2011). Auch bei Zhao und Kollegen wird die Funktion der Facebook-*Timeline* als Selbstbild nach Außen, welches dann auch verwaltet wird, angesprochen (Zhao et al., 2013). Der Umstand, dass von den Probanden neben Twitter auch andere Social-Media-Anwendungen genutzt werden, die noch viel stärker als Archiv oder Werkzeug zur Selbstdarstellung wahrgenommen werden, kann auch die Wahrnehmung bzw. Charakterisierung Twitters beeinflussen.

Viele Kommentare unterstützen die Wahrnehmung der Funktion des Favorisierens als *bookmarking*. Favorisieren als Lesezeichen-Funktion ist das Konzept, das am stärksten mit der Fragestellung nach dem Wiederfindensverhalten auf Twitter in Beziehung steht. Das Aufbewahrensverhalten ist ein zentraler Bestandteil des PIM und auch für das Wiederfinden per se. Durch Favorisieren werden Tweets in einer separaten Liste abseits der *Timeline* gespeichert, um einen erneuten Zugang zu beschleunigen, also das Wiederfinden von Tweets einfacher zu gestalten. Obwohl Nutzer angeben, Tweets zu favorisieren, um sie in der Favoritenliste aufzubewahren und dann dort zu einem späteren Zeitpunkt aufzurufen, wird die Favoritenliste laut den Umfrageergebnissen nicht sehr häufig besucht (vgl. Abbildung 5.13 in Abschnitt 5.6.2). Darüber hinaus deuten einige Nutzerkommentare darauf hin, dass Tweets aus der Favoritenliste gelöscht werden (Entfavorisieren), nachdem sie erneut aufgerufen wurden. Die Freitextfrage 2.10 *Can you name reasons that lead you to unfavourite Tweets* wurde zwar nicht systematisch analysiert, wie dies bei der Frage nach den Favorisierungsgründen der Fall war, jedoch kann festgestellt werden, dass Entfavorisieren, aufgrund eines nicht mehr benötigten Bookmarks, ein häufig genannter Grund war. Kommentare wie „I would unfavorite if I had no reason to keep them bookmarked any longer“ oder „I have read the linked article“ belegen, dass favorisierte Tweets den Wert verlieren, wenn die Nachfolgeaktion abgeschlossen ist. Die Umfragedaten zeigen jedoch auch hier, dass das Entfavorisieren eine sehr seltene Aktion ist. Anhand der Begründungen ist zu erkennen, dass das Wiederfinden nicht das einzige Ziel ist, das Nutzer mit dem Erstellen von Lesezeichen auf einem Tweet verfolgen. So wird es auch als Archiv genutzt, um eine Liste der schönsten Tweets zu kuratieren. Dass Nutzer bestimmte Features ei-

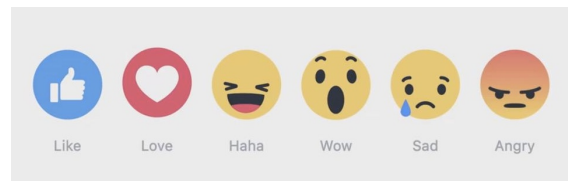
ner Social-Media-Plattform als Life-logging-Werkzeug verwenden, um ein persönliches Archiv aus Nachrichten, Kommentaren oder Bildern aufzubauen, an die sie sich gerne erinnern bzw. ihre eigenen digitalen Spuren zu bewahren, wurde bereits in früheren Arbeiten zu Facebook deutlich (Zhao et al., 2013). Dies steht wiederum in Diskrepanz zur öffentlichen Zugänglichkeit der Favoritenliste und der vermeintlich privaten Form der Zustimmung.

Man erkennt: die Favoriten-Funktion wird zweckentfremdet. Oder wie Zhao und Kollegen es formulieren: „But in practice, people do use Social Media for multiple purposes and designs should respect that“ (Zhao et al., 2013, S.9). Nutzer haben viele verschiedene Bedürfnisse aber wenig Funktionalität zur Hand, so dass der Favoriten-Button für eine Vielzahl von unterschiedlichen Absichten verwendet wird. Eine Frage, die sich hier stellt ist, ob die Einführung von weiteren Buttons eine Erleichterung für die Nutzer bringen würde, da sich Nutzerabsichten auf mehr Einzelbuttons verteilen würden. Eine mögliche Lösung für diesen Konflikt könnte die zusätzliche Einführung einer Speicher-Funktion sein, die tatsächlich wie eine Lesezeichenfunktion funktioniert und Tweets in einer privaten Liste speichert und damit diese Funktionalität von den Kommunikationsfunktionen, dem *liking* und dem Zeigen von Anerkennung, konzeptuell trennt. Diese Änderung würde schließlich auch das Wiederfinden von Tweets erleichtern. Folglich könnte eine Funktion implementiert werden, die Tweets zu den Favoriten hinzufügt und dort speichert, wohingegen es eine zusätzliche Funktion geben könnte, die lediglich zu Kommunikationszwecken dient und betroffene Tweets nicht speichert. Auch für den Zweck des Erstellens einer Liste von erinnerswerten persönlichen Tweets, bzw. der Nutzung als persönliches Archiv wäre eine private Like-Funktion von Vorteil. Den Vorschlag eines privaten *like*, den nur der Nutzer selbst sieht, ist ein Vorschlag, der auch von Zhao und Lindley in Bezug auf die Like-Funktion von Facebook gemacht wird, um die Funktionalität von Facebook als persönliches Archiv zu verbessern (Zhao & Lindley, 2014).

Schließlich muss in Erinnerung gerufen werden, dass knapp 35% der Studienteilnehmer angeben, die Favoriten-Funktion nicht zu kennen. Dies kann dadurch begründet sein, dass Twitter seitens dieser Teilnehmer zu selten genutzt wird, oder die Favoriten-Funktion tatsächlich zu wenig selbst-erklärend ist und ihr Nutzen für diese Teilnehmer nicht offensichtlich genug ist. Insofern könnte zusätzlich auch darüber nachgedacht werden, die Funktion präsenter zu machen und ihr in diesem Zuge mehr Eindeutigkeit zu verleihen, wie bereits oben beschrieben. Fraglich ist auch, wie die Weiterentwicklung der Favoriten-Funktion die Nutzungspraktiken der Funktion beeinflusst hat. Obwohl Favorisieren bereits eine Funktion der ersten Stunde ist, wurde die Anzeige wie viele Favorisierungen ein Tweet erhalten hat, erst zu einem späteren Zeitpunkt eingeführt. Dies führte in der Community schließlich auch zur Entwicklung der Plattform Favstar, die seit jeher die Anzeige der erhaltenen Favoriten ermöglichte und die in Kode [B3.1] beschriebenen nicht-kommerziellen Wettbewerbe forcierte. Somit nehmen auch Drittanbieter und deren Plattformen Einfluss darauf, wie Nutzer die Funktion wahrnehmen und verwenden (Paßmann & Gerlitz, 2014).

Bleibt abschließend die Frage zu diskutieren, ob sich die gewonnenen Erkenntnisse auf andere Plattformen übertragen lassen und somit generalisierbar sind. Es stellt sich die Frage, welche Konsequenzen das Wissen über die Nutzung der Favoriten-Funktion von Twitter auf die Wahrnehmung der Nutzung von vergleichbaren Features besitzt, die eine ähnliche Funktionalität besitzen wie der *like* bei Facebook, *+1* bei Google Plus (G+) oder *hearts* auf Instagram, Tumblr und Pinterest. Facebook und G+ bieten dem Nutzer viel mehr Möglichkeiten die Funktionalität einzusetzen, da dort nicht nur Beiträge von Nutzern geliked werden können sondern auch Kommentare, Personen, Orte, Ereignisse und Themen. Diese werden anschließend auf der Profilseite des Nutzers angezeigt, ähnlich wie die favorisierten Tweets in der Favoritenliste des Nutzers auf Twitter. Obwohl es nur ein Feature gibt, ist es sehr wahrscheinlich, dass schon allein die Mehrzahl an Objekten, die geliked werden können, das Nutzungsverhalten der Funktionalität im allgemeinen verändert. Facebook hat auch bereits darauf reagiert, dass Nutzer viele unterschiedliche Bedürfnisse mit einem Button

zu kommunizieren versuchen und den Like-Button ausdifferenziert, so dass Facebook-Nutzer mittlerweile sechs unterschiedliche Icons, welche sechs verschiedene Gefühlszustände repräsentieren, zur Verfügung haben (Krug, 2016). Abbildung 5.14 zeigt die sechs Icons, welche im Februar 2016 global eingeführt wurden und als *reactions* bezeichnet werden.



**Abbildung 5.14.:** Die sechs Reaktionsmöglichkeiten, die das klassische Like abgelöst haben

(Screenshot von: <http://newsroom.fb.com/news/2016/02/reactions-now-available-globally/>)

Bei Tumblr und Pinterest — wie bei Twitter — können Nutzer nur die jeweiligen Posts als einzelne Objekte favorisieren bzw. liken. Möglicherweise besteht somit das Bedürfnis nach mehr Funktionalität, um die aus den Nutzerkommentaren herausgearbeiteten Motive auch alle zu unterstützen. Dass der Like-Funktionalität per se aber in den letzten Jahren durchaus zu wenig Aufmerksamkeit zugekommen ist und dort eine gewisse Forschungslücke bestand, zeigt eine vermehrte Anzahl an Studien, die sich mit vergleichbaren Features auf anderen Plattformen auseinandersetzt und zum Teil auch die hier vorgestellten Analysen zitieren. Die Studie *One Click, Many Meanings* von Hayes, Carr und Wohn untersucht die Nutzung der Like-Funktion über mehrere Plattformen hinweg mit Hilfe einer Fokus-Gruppen-Studie (Hayes et al., 2016). Lee, Hansen und Lee führen ebenfalls eine Umfrage zur Nutzung der Like-Funktion auf Facebook durch (Lee et al., 2016). Levordashka, Utz und Ambros untersuchen das Spektrum an Beweggründen für das Klicken des Like-Button auf Facebook (Levordashka et al., 2016). Mitarbeiter von Facebook selbst analysieren Motive für die Nutzung der Like-Funktion. Jang, Han und Lee untersuchen in einer quantitativen Studie das Like-Verhalten von Nutzern auf Instagram, wobei sie feststellen, dass Likes auf dieser Plattform selten auf Gegenseitigkeit beruht (Jang et al., 2015). Motive für die Nutzung der Like-Funktion werden nicht analysiert, jedoch ermitteln sie fünf Faktoren, die dazu beitragen, dass Nutzer viele Likes erhalten (Jang et al., 2015).

Die hier beschriebene Studie zur Nutzung der Favoriten-Funktion auf Twitter kann durchaus als Pionierarbeit betrachtet werden, da vergleichbare Arbeiten erst in den Folgejahren veröffentlicht wurden.

## 5.7. Quantitative Ergebnisse zu Aufbewahrens- und Wiederfindensstrategien

Abschnitt 5.6.3.2 hat offengelegt, dass Nutzer die Favoriten-Funktion als Lesezeichenfunktion nutzen, um bedeutsame Tweets in einer separaten Liste zu speichern und somit die Rückkehr zu diesen Tweets zu garantieren und potentiell zu beschleunigen. Fraglich ist dabei, wie die Nutzung der Favoriten-Funktion im Verhältnis zu anderen Aufbewahrensstrategien steht, wie häufig Nutzer generell Tweets aufbewahren, wiederfinden und welche Strategien sie einsetzen. Da Aufbewahren

und Wiederfinden komplementäre Aktivitäten sind und bestimmte Wiederfindensstrategien das Aufbewahren, also *keeping behaviour*, voraussetzen, gilt es zunächst, die Häufigkeit und Strategien des Aufbewahrens zu betrachten.

In den Abschnitten 5.7.2 bis 5.7.7 tritt das wiederkehrende Problem auf, dass Unterschiede zwischen Gruppen ausgemacht werden müssen. z. B. soll die Frage geklärt werden, ob unter den Probanden signifikante Unterschiede in der Nutzung von Wiederfindensstrategien bestehen, wenn diese häufiger bzw. weniger häufig Tweets aufbewahren. Zur Signifikanzprüfung stehen hier der parametrische T-Test — unter Voraussetzung der Parametererfüllung wie Normalverteilung und Varianzhomogenität — oder der parameterfreie Wilcoxon-Test, wenn die jeweiligen Voraussetzungen nicht erfüllt sind, zur Verfügung. Um jedoch ein einheitliches Verfahren bei allen Signifikanzprüfungen anzuwenden und aufgrund der Vorteile des nicht-parametrischen Bootstraps gegenüber Standardverfahren, wurde dieser bei allen Prüfungen vorgezogen. Bei diesem Verfahren wird anstelle eines Signifikanztests ein *indirekter* Hypothesentest mithilfe von Konfidenzintervallen, welche auf der Basis von Bootstrap-Stichproben errechnet werden, verwendet. Die Signifikanzprüfung mithilfe des nicht-parametrischen Bootstraps wird im Folgenden vorgestellt.

### 5.7.1. Signifikanzprüfung mithilfe des nicht-parametrischen Bootstraps

In der Testfamilie der parameterfreien Signifikanztests bietet sich der für unabhängige Stichproben geeignete Rangsummentest nach Wilcoxon an, von dessen Verwendung jedoch Abstand genommen wird. Der Verzicht auf Signifikanztests zu Gunsten von Konfidenzintervallen ist eine Diskussion, die seit mehreren Jahrzehnten in medizinischen und psychologischen Fachzeitschriften geführt wird, aber außerhalb des Fokus dieser Arbeit steht. Für eine vertiefte Auseinandersetzung mit dieser Thematik siehe z. B. Brandstätter (Brandstätter, 1999).

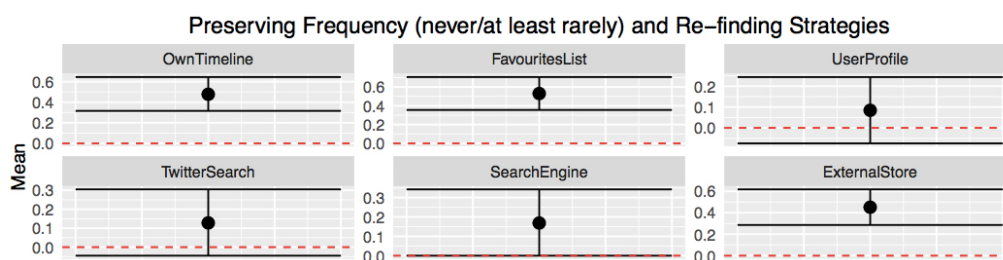
Ein Konfidenzintervall ist ein Wertebereich für einen statistischen Populationsparameter (z. B. Mittelwert, Median oder das *odds ratio*), der als zusätzliche Information zum Kennwert über die Unsicherheit eines Wertes berichtet. Es bildet den Gegenpart zur Irrtumswahrscheinlichkeit, wird häufig auch als Sicherheitswahrscheinlichkeit bezeichnet und kann somit als Maß der Genauigkeit verstanden werden. Pragmatisch lässt sich das Konfidenzintervall — ein Signifikanzniveau von 5% vorausgesetzt — etwa wie folgt charakterisieren: Man kann mit einer Sicherheit von 95% sagen, dass bei einer beliebig häufigen Replikation einer Studie, der wahre Wert eines Populationsparameters in dem Bereich des Konfidenzintervalls liegt. Konfidenzintervalle können dazu genutzt werden, um die Nullhypothese *indirekt* zu testen, bzw. um Rückschlüsse auf die statistische Signifikanz zu ziehen, da, wie bei allen statistischen Tests, auf der Basis einer Stichprobe auf die Grundgesamtheit geschlossen werden kann. Dies wird realisiert, indem geprüft wird, ob die Konfidenzintervalle des entsprechenden Parameters den Wert des Null-Effekts enthalten oder nicht. Bei der Differenz zweier Mittelwerte wäre dies die 0, bei einem *odds ratio* die 1 (Bender & Lange, 2007). Aufgrund des Zentralen Grenzwertsatzes können Konfidenzintervalle bei einer Normalverteilung des interessierenden Parameters über die hierfür bekannten z-Werte berechnet werden (bei 95% Konfidenzintervall  $z = 1.96$ ). Liegt keine Normalverteilung vor bzw. ist die Verteilungsfunktion eines Parameters in der Stichprobe unbekannt, muss eine andere Methode zur Berechnung der Konfidenzintervalle gefunden werden. Eine Lösung bietet die Bestimmung von Konfidenzintervallen mithilfe von Bootstrap-Stichproben.

Das *bootstrapping* gehört zur Familie der Resampling-Verfahren und ist verwandt mit anderen Permutationsmethoden wie der *cross validation* oder dem *jackknife*. Sie sind alle dadurch charakterisiert, dass durch häufiges Ziehen *mit/ohne Zurücklegen* ein oder mehrere neue Stichproben (*Samples*) aus der Basisstichprobe erzeugt werden. Als Ausgangspunkt für das Bootstrap-Verfahren dient die konkret vorliegende Stichprobe, die als Approximation der Grundgesamtheit gesehen



wird und somit ein repräsentatives Abbild darstellt auf dessen Basis die Grundgesamtheit simuliert werden kann (Shikano, 2010). Die konkrete Stichprobe nennt sich auch Basisstichprobe und dient als Grundlage für den Bootstrap, der nun darin besteht, dass  $B$  Stichproben im Umfang der Größe  $n$  mit Zurücklegen aus der Basisstichprobe gezogen werden. Die gezogenen Stichproben werden dann als Bootstrap-Stichproben bezeichnet (Holling & Gediga, 2015). Für jede dieser Bootstrap-Stichproben wird der sogenannte „interessierende Kennwert“  $\theta$  berechnet (Shikano, 2010, S.193). Dies können unterschiedliche statistische Kennwerte sein z. B. der Median, Differenz von Mittelwerten oder auch Regressionskoeffizienten. Man spricht hier oft auch vom sogenannten Plug-in-Prinzip, da der Populationswert des interessierenden Parameters und die Standardabweichung über den Stichprobenwert und die Stichprobenstandardabweichung geschätzt werden. Ergo dienen die Stichprobenwerte als Platzhalterwerte. Im vorliegenden Fall wird zum Vergleich zweier Stichproben die Differenz derer Mittelwerte berechnet, wie bei Chernick (Michael R. Chernick, 2008, S.67-71), Wood (Wood, 2005, S.462ff.) oder Hesterberg (Hesterberg et al., 2010, S.14-18ff.) beschrieben. Auf der Basis der Stichprobenverteilung für den jeweiligen Parameter  $\theta$  werden die Bootstrap-Konfidenzintervalle bestimmt. Ein einfaches Verfahren stellt hier die Verwendung des Bootstrap-Perzentil-Intervalls (Perzentil-Methode) für  $\theta$  dar, bei dem die Quantile der Bootstrap-Verteilung auf der Basis des Alphaniveaus zur Bestimmung der Konfidenzintervalle verwendet werden. Bei einem Wert von  $\alpha = 0,05$  werden das 0,025- und das 0,975-Quantil genutzt und als Unter- und Obergrenze für das Konfidenzintervall festgelegt. Ein signifikantes Ergebnis liegt vor, wenn die Konfidenzintervalle den Null-Effekt nicht enthalten, also die Fehlerbalken, welche die Werte der Konfidenzintervalle repräsentieren, die Null-Werte nicht berühren.

Bei den folgenden Gruppenvergleichen werden pro Gruppe  $B = 10000$  Stichproben der Größe  $n$  gezogen, wobei sich die Anzahl  $n$  immer auf die Anzahl an Beobachtungen der größeren Gruppe bezieht. Bei jeder Ziehung wird für beide Gruppen der Mittelwert berechnet und anschließend subtrahiert. Schließlich werden die Konfidenzintervalle des subtrahierten Wertes gebildet. Dies kann anschließend visualisiert werden. Die so entstehenden Grafiken zeigen, ob der Nulleffekt existiert — also die Verteilungen nicht signifikant unterschiedlich sind — oder nicht. Abbildung 5.15 zeigt exemplarisch die erstellten Abbildungen für die Frage, ob ein signifikanter Unterschied in der Zustimmung zur Nutzung bestimmter Wiederfindensstrategien besteht, wenn Nutzer angeben nie oder zumindest selten Tweets aufzubewahren. Wenn der Fehlerbalken die rote gestrichelte Linie berührt, deutet dies auf einen Nulleffekt hin, so dass nicht von einem signifikanten Unterschied ausgegangen werden kann. Erfolgt keine Überschneidung des Fehlerbalkens mit der Nulllinie, so kann von einem signifikantem Unterschied ausgegangen werden.

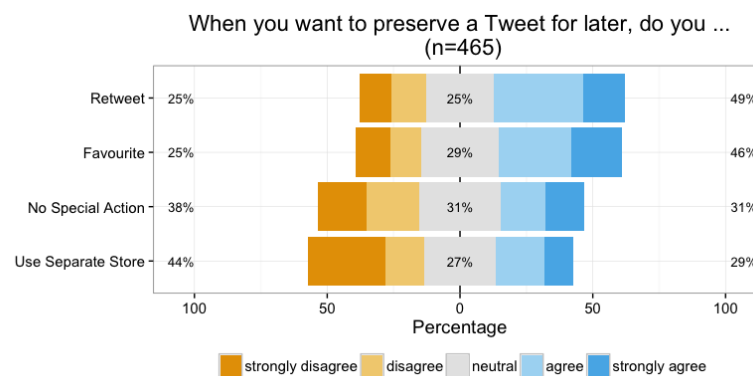


**Abbildung 5.15.:** Signifikanzprüfung durch visuelle Begutachtung der Konfidenzintervalle

Die erstellten Plots aller Signifikanztests sind dem Anhang A.2 zu entnehmen.

### 5.7.2. Aufbewahrenshäufigkeit und -strategien

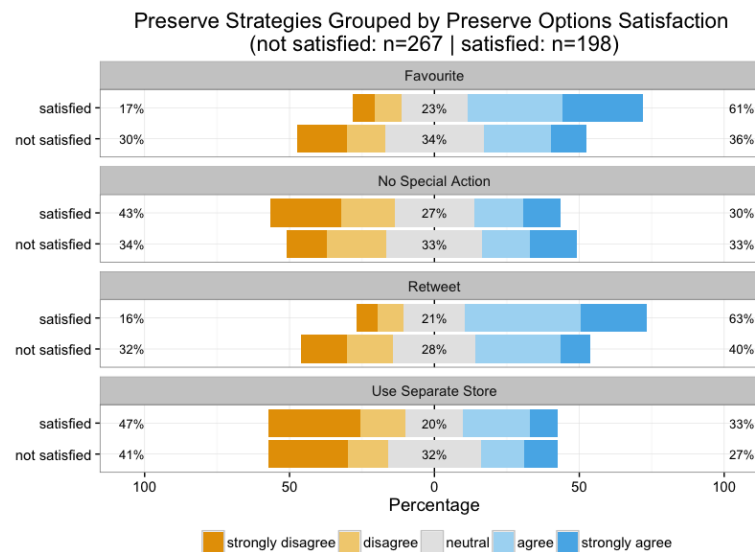
Die Frage 3.1 *How often do you come across a Tweet, which you consider worth preserving for later?* beantworteten nur 22.7% der Teilnehmer mit nie (*never*), woraus erkenntlich wird, dass über drei Viertel der Probanden (77,3%) schon einmal das Bedürfnis verspürt haben, einen Tweet für einen späteren Zeitpunkt aufzubewahren. Für 44% der Teilnehmer ist dies jedoch ein seltenes Ereignis. Nur 2,5% besitzen dieses Bedürfnis täglich. Eine mögliche Erklärung, weshalb Nutzer den Eindruck haben, nur selten in die Situation zu kommen, Tweets aufbewahren zu müssen, könnte das Vertrauen in die eigene Fähigkeit sein, Tweets auf einfache Weise wiederfinden zu können bzw. das Vertrauen in die gegebenen Twitter-Optionen, die das Zurückkehren zu Tweets erleichtern. Begründungen für diese Vermutung lassen sich in unterschiedlichen Wiederfindensstrategien von Nutzern finden, respektive in unterschiedlichen Methoden für Nutzer, die angeben, dass die Optionen, die Twitter bietet, um Tweets wiederzufinden, gut oder sehr gut sind. Diese Aspekte werden später noch im Detail analysiert.



**Abbildung 5.16.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Aufbewahrensstrategien

Betrachtet man die Strategien, die Nutzer (vgl. Abbildung 5.16) angeben, dass sie diese zum Aufbewahren von Tweets verwenden, wird deutlich, dass beinahe die Hälfte (49% und 46%) aller Probanden der Aussage zustimmen, die Funktionen Weiterleiten (RT) und Favorisieren zum Aufbewahren zu nutzen. 29% der Probanden geben an, Tweets in einem externen Speicherort zu verwahren (z. B. via Lesezeichen im Browser), wobei jedoch 44% der Studienteilnehmer der Aussage, diese Handlung vorzunehmen, widersprechen. Dies zeigt eine starke Präferenz seitens der Nutzer zur Verwendung von internen Twitter-Funktionen zum Speichern und Aufbewahren von Tweets. Nur 31% der Studienteilnehmer stimmen der Aussage zu, keine weitere Aktion vorzunehmen, so dass davon ausgegangen werden kann, dass Nutzer, wenn sie Tweets als aufbewahrenswert erachten, mit einer entsprechenden Reaktion darauf reagieren. Dabei stellt sich die Frage, ob Probanden, die die Twitter-Optionen zum Aufbewahren von Tweets als schlecht oder sehr schlecht bewerten, eher dazu tendieren, anders zu handeln als Nutzer, die mit den Funktionen zufrieden sind. Zu diesem Zweck erfolgt eine Gruppierung entlang der Aussagen zur Verwendung von Aufbewahrensstrategien in Nutzer, die die Aufbewahrensfunktionen als schlecht bzw. gut charakterisieren.

Abbildung 5.17 zeigt den Widerspruch bzw. die Zustimmung der Studienteilnehmer zur Verwendung der jeweiligen Aufbewahrensstrategien gruppiert nach Nutzern, die die twitterinternen Auf-

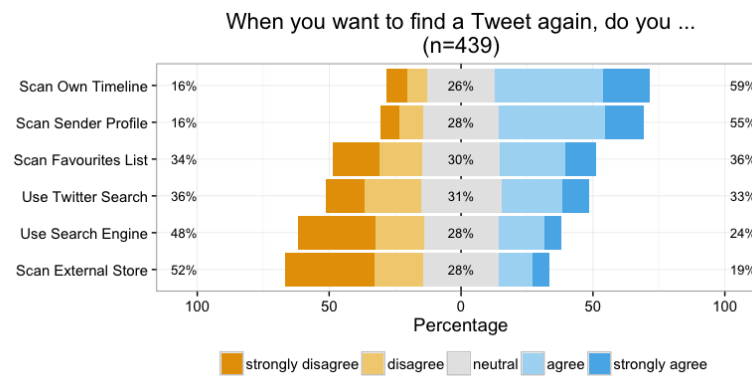


**Abbildung 5.17.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Aufbewahrensstrategien gruppiert nach Bewertung der Aufbewahrensfunktionen

bewahrensstrategien als gut oder sehr gut charakterisieren und Nutzern, die die Funktionen als nicht zufriedenstellend bewerten. Obwohl die Mehrheit der Probanden (267) die Aufbewahrensfunktionen als nicht zufriedenstellend charakterisieren, ist die Tendenz zur Nutzung von Weiterleiten und Favorisieren zur Aufbewahrung in der Gruppe der zufriedenen Nutzer noch stärker zu erkennen. Eine Prüfung auf Signifikanz unter Verwendung des nicht-parametrischen Bootstrap-Verfahrens zeigt, dass Nutzer, die die Funktionen als zufriedenstellend charakterisieren, auch signifikant stärker der Aussage zustimmen, diese zu nutzen. Interessanterweise verkehrt sich dies bei der Aussage nach der Verwendung keiner speziellen Interaktion ins Gegenteil: Studienteilnehmer, die die Aufbewahrensfunktionen auf der Plattform als zufriedenstellend empfinden, widersprechen der Aussage, keine Aktionen vorzunehmen (Antwortmöglichkeit: *No Special Action*) signifikant stärker, wodurch sich ein rundes Bild ergibt. Für die Verwendung eines externen Speichers lassen sich keine signifikanten Unterschiede zwischen den Gruppen feststellen.

### 5.7.3. Wiederfindenshäufigkeit und -strategien

Fast drei Viertel aller Probanden (72,4%) berichten schon einmal das Problem gehabt zu haben, zu bereits gesehenen Tweets zurückkehren zu wollen. 26,6% der Probanden geben an, dass sie noch nie mit diesem Problem konfrontiert waren. Somit ist die Zahl der Probanden, die Tweets wiederfinden geringer als die der Teilnehmer, die Tweets aufbewahren. Eine mögliche Erklärung bietet die Nutzung der Favoriten-Funktion/Favoritenliste zur Sammlung erinnerenswürdiger Tweets (Kode [B 2.1] der Favorisierungsgründe). Nutzer entscheiden sich somit häufiger, Tweets in ihre Kollektion der besonderen Tweets aufzunehmen, ohne tatsächlich jedes Mal das Bedürfnis nach einem unmittelbaren erneuten Aufruf zu besitzen. Für 43,9% ist das Wiederfinden von Tweets eine Aktion, die eher selten vorkommt. Zumindest 8,4% der Teilnehmer geben an, wöchentlich oder noch häufiger, Tweets wiederfinden zu wollen. Dabei stellt sich die Frage, welche Strategien sie verfolgen bzw. Methoden sie nutzen, um Tweets wiederzufinden.



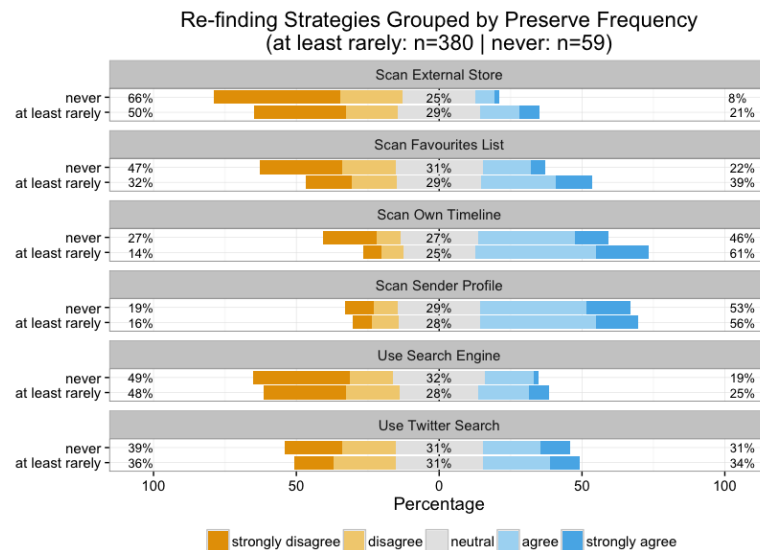
**Abbildung 5.18.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien

Abbildung 5.18 zeigt das Umfrageergebnis zur Frage 4.4 *When you want to find a Tweet again, do you....* Die Zustimmung zu den Aussagen repräsentieren mögliche Wiederfindensstrategien. Die Visualisierung der Likert-Skalen macht deutlich, dass knapp 60% aller Teilnehmer der Aussage zustimmen, Tweets in der eigenen *Timeline* zu lokalisieren (Strategie: *Scan Own Timeline*. 55% der Probanden stimmen der Aussage zu, das Profil des Tweet-Senders nach dem jeweiligen Ziel-Tweet zu durchsuchen (Strategie *Scan Sender Profile*). Beide Wiederfindensstrategien können in Abhängigkeit von bestimmten Faktoren durchaus komplexe Handlungen sein. Faktoren, die zur Komplexität beitragen, sind: die Zeit, die seit der ersten Begegnung mit einem Tweet vergangen ist und die Anzahl der gefolgten Nutzer respektive die Tweetfrequenz des Senders, ergo die Überladung der jeweiligen *Timeline*. Sind beide Parameter sehr hoch, ist ein erheblicher Scroll-Aufwand erforderlich, um den Zieltweet zu finden. An dritter Stelle findet sich mit 36% Zustimmung das Lokalisieren des Tweets in der Favoritenliste und damit eine Wiederfindensstrategie, die eine Aufbewahrensstrategie voraussetzt. Nur ein Drittel (33%) der Probanden stimmt der Aussage zu, die twitterinterne Suchfunktion zu nutzen, um Tweets wiederzufinden. Immerhin 24% geben an, dass die Verwendung einer klassischen Suchmaschine wie Google oder Bing einen möglichen Weg darstellt, um Tweets erneut aufzufinden. Die Verwendung eines externen Speicherortes, der ebenfalls das vorherige Verwahren eines Tweets voraussetzt, wird nur von 19% als mögliche Wiederfindensstrategie genannt. Die Tendenz, dass externe Speicherorte eine untergeordnete Rolle für die Verwaltung von Tweets spielen, bestätigt sich somit.

In den vorherigen Abschnitten wurde bereits deutlich, dass Probanden häufiger den Eindruck haben, Tweets zu lesen, die sie als aufbewahrenswert charakterisieren, als tatsächlich das Bedürfnis zu haben, diese wiederzufinden. Möglich ist dabei auch, dass die Häufigkeit des Aufbewahrens einen Einfluss darauf hat, wie Tweets wiedergefunden werden, also welche Strategien für unterschiedliche Nutzergruppen (nie aufbewahren vs. häufig aufbewahren) eher in Frage kommen. Neben der Aufbewahrensfrequenz können auch andere Faktoren die Tendenz, bestimmte Wiederfindensstrategien zu verwenden, beeinflussen. Insgesamt werden folgende Einflussfaktoren diskutiert und in den folgenden Kapiteln näher ausgeführt:

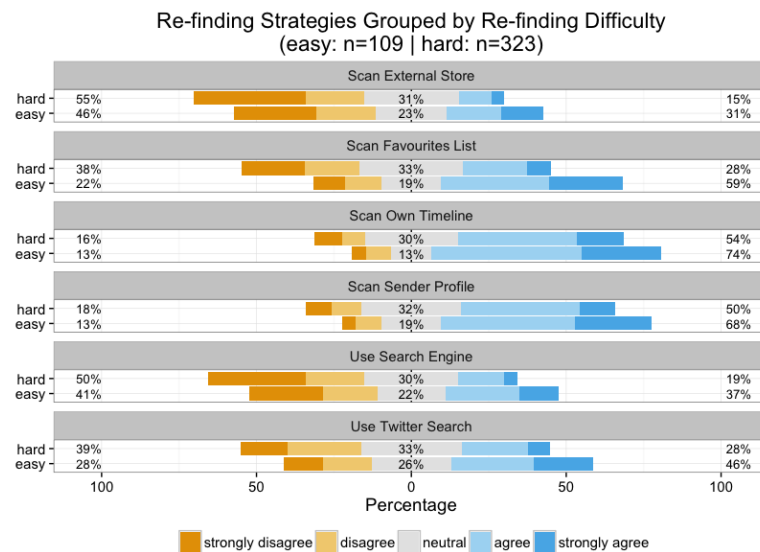
- Aufbewahrenshäufigkeit
- Schwierigkeit des Wiederfindens
- Bewertung von Re-finding-Optionen
- Frustrationserfahrung

### 5.7.4. Aufbewahrenshäufigkeit und Wiederfindensstrategien



**Abbildung 5.19.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung von Wiederfindensstrategien gruppiert nach Häufigkeit des Aufbewahrens

Abbildung 5.19 zeigt die Aufteilung der Wiederfindensstrategien in die Gruppe der Probanden, die aussagt, nie Tweets aufzubewahren bzw. die Gruppe der Probanden, die angibt, dies zumindest selten oder häufiger zu tun. Bei beiden Gruppen ist die Zustimmung zum Wiederauffinden von Tweets in der eigenen *Timeline* (Gruppe *never*: 46% vs. Gruppe *at least rarely*: 61%) und in den gesendeten Tweets von Nutzern (Gruppe *never*: 53% vs. Gruppe *at least rarely*: 56%) sehr ausgeprägt. Die Gruppe der Probanden, die angibt, nie aufzubewahren, ist zwar mit 59 Personen eher klein, dennoch ist dort eine stärkere Abneigung gegen Wiederfindensstrategien zu erkennen, die ein vorheriges Aufbewahren erfordern, wie die Verwendung eines externen Speicherorts (66%) und der Favoritenliste (47%). Erstaunlich ist, dass bei dieser Gruppe der Widerspruch zur Aussage, Suchmaschinen zum Wiederfinden zu verwenden, sehr hoch ist. Mithilfe des Bootstrap-Verfahrens können innerhalb der Gruppen signifikante Unterschiede in der Zustimmung zur Verwendung von Wiederfindensstrategien festgestellt werden. Die Gruppe der Nutzer, die zumindest selten Tweets aufbewahrt, stimmt den Aussagen über die eigene *Timeline*, die Favoritenliste und externe Speicher wiederzufinden signifikant stärker zu, als Probanden, die nie Tweets aufbewahren. Bei den anderen Wiederfindensstrategien, wie der Suche in den gesendeten Tweets von Nutzer, der Nutzung der Twitter-Suche oder einer externen Suchmaschine, lassen sich keine signifikanten Unterschiede in der Zustimmung ausmachen. Es lässt sich also feststellen, dass Probanden, die angeben, Wiederfindensstrategien zu nutzen, die ein Aufbewahren voraussetzen, diese Strategien auch konsequent und häufiger nutzen.



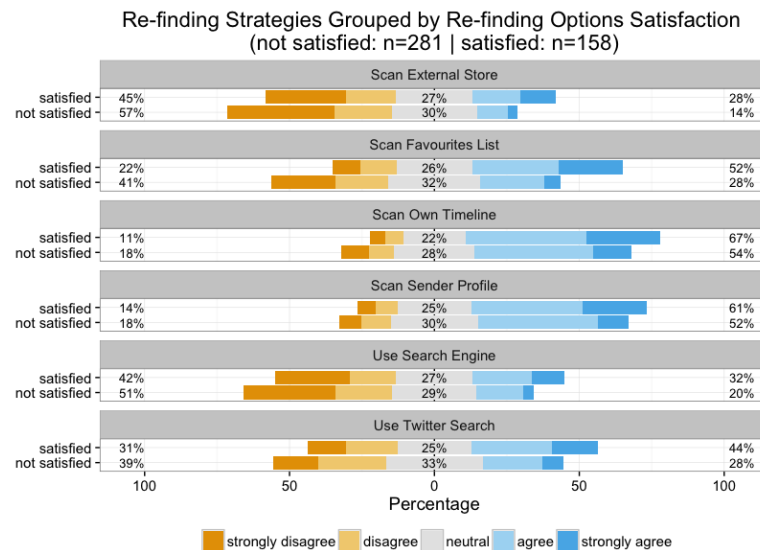
**Abbildung 5.20.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien gruppiert nach Einschätzung des Wiederfindensprozesses

### 5.7.5. Schwierigkeit des Wiederfindens und Wiederfindensstrategien

Frage 4.8 befragte die Teilnehmer, wie stark sie der Aussage zustimmen, dass Wiederfinden von Tweets auf Twitter einfach sei (Frage 4.8 *How strong do you agree with these statements? Re-finding on Twitter is easy*). Bei der Beantwortung der Frage schlägt sich eine in der Literatur bekannte Antworttendenz, nämlich die Tendenz zur Mitte, deutlich durch, da 43,1% die Frage mit *neutral* beantworten (Kallus, 2010, S.55). Der Anteil an Umfrageteilnehmern, die dieser Frage widersprechen, liegt bei 31,9% und damit um etwa 7% höher als der Anteil der Probanden, die aussagen, dass Wiederfinden einfach ist. Im Folgenden wird untersucht, ob Nutzer, die das Wiederfinden auf Twitter als einfach charakterisieren, andere Wiederfindensstrategien verfolgen als Nutzer, die dieser Aussage widersprechen.

Abbildung 5.20 zeigt die Zustimmung zur Verwendung der Wiederfindensstrategien gruppiert nach Nutzergruppen, die Wiederfinden als einfache Aufgabe charakterisieren bzw. dieser Aussage nicht zustimmen. Man kann erkennen, dass das Durchsuchen der eigenen *Timeline* sowie das Durchsuchen der gesendeten Tweets von anderen Nutzern bei beiden Gruppen die dominanten Strategien darstellen. In der Gruppe der Teilnehmer, die Re-finding als einfach charakterisieren, ergeben sich sogar sehr hohe Werte (68% und 74%) an Zustimmung für die Nutzung dieser Strategien. Zudem ist zu erkennen, dass bei allen Strategien, die Gruppe der Nutzer, die Wiederfinden als schwierige Aufgabe charakterisieren, der Widerspruch zur Verwendung der Strategien deutlich höher ausfällt, als bei der Gruppe, die Re-finding als einfach charakterisiert. Dieser Unterschied ist für alle Strategien signifikant. Probanden, die Wiederfinden als einfach charakterisieren, stimmen den Aussagen der Nutzung aller Strategien signifikant stärker zu. Dies lässt den Schluss zu, dass das Wiederfinden dann als einfach charakterisiert wird, wenn auch alle oder zumindest mehrere Wiederfindensstrategien genutzt werden.

### 5.7.6. Bewertung der Twitteroptionen und Wiederfindensstrategien



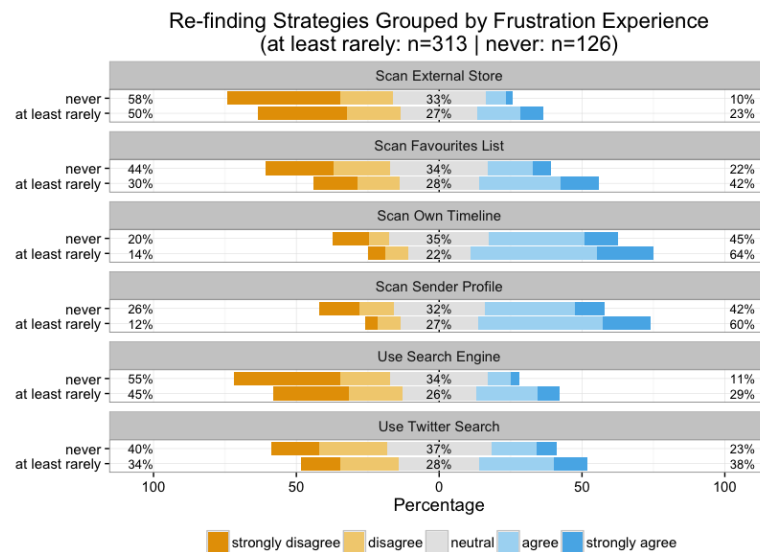
**Abbildung 5.21.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien gruppiert nach Bewertung der Twitter-Optionen für Wiederfinden

Bei der Zustimmung zur Frage 5.1 *How would you rate Twitter's current facilities/options for: Finding previously seen Tweets again*, ist ebenfalls eine starke Tendenz zur Mitte zu erkennen (47,2%). 32,45% bewerten die Optionen für das Wiederfinden auf Twitter als gut oder sehr gut. Nur knapp über 20% bewerten die Optionen als schlecht oder sehr schlecht. Ein zum vorherigen Abschnitt vergleichbares Bild ergibt sich auch bei der Bewertung der Twitter-Optionen und der Zustimmung zur Verwendung bestimmter Strategien (vgl. Abbildung 5.21). Die Strategien *Scan Own Timeline* sowie *Scan Sender Profile* dominieren in beiden Gruppen. Des Weiteren lässt sich die generelle Tendenz bestätigen, dass Nutzer, die stärker der Verwendung aller Strategien widersprechen, auch unzufriedener sind mit den Wiederfindensoptionen, die Twitter bietet. Der nicht-parametrische Bootstrap zeigt signifikante Unterschiede zwischen der Gruppe der zufriedenen und unzufriedenen Nutzern für alle Wiederfindensstrategien. Analog zum vorherigen Abschnitt lässt dies den Schluß zu: Nutzer, die alle Strategien zum Wiederfinden verwenden, sind auch zufriedener mit den twitterinternen Optionen für Re-finding.

### 5.7.7. Frustrationserfahrung und Wiederfindensstrategien

Im Folgenden wird untersucht, ob Nutzer, die bereits einmal ein frustrierendes Wiederfindenserlebnis erfahren haben (frustrierend kann interpretiert werden als erfolgloser oder sehr lange andauernder Wiederfindensprozess), andere Strategien verwenden als Nutzer, die noch nie ein Negativerlebnis in dieser Hinsicht hatten. Frage 5.2 der Twitter-Umfrage befragte Nutzer nach der Häufigkeit, mit der sie ein frustrierendes Erlebnis beim Re-finding-Prozess erfahren haben. 58,3% berichten, dass sie schon einmal ein frustrierendes Erlebnis bei der Suche nach einem be-





**Abbildung 5.22.:** Stärke der Ablehnung bzw. Zustimmung zur Verwendung der genannten Wiederfindensstrategien gruppiert nach Frustrationserfahrung

reits gesehenen Tweet gehabt haben. Für 20% der Teilnehmer passiert dies regelmäßig. 37,9% der Teilnehmer sind aber nur selten davon betroffen. Nichtsdestotrotz zeichnet sich ein Bild, dass Tweet-Re-finding ein frustrierendes Sucherlebnis sein kann.

Abbildung 5.22 zeigt in Analogie zu den vorherigen Analysen die Zustimmung zu Wiederfindensstrategien gruppiert nach Probandengruppen mit frustrierendem und ohne frustrierendes Wiederfindenserlebnis. Auch hier ergibt sich ein ähnliches Bild wie in den beiden Abschnitten zuvor: Nutzer, die noch nie frustriert waren, stimmen signifikant stärker den Aussagen zu, jede Wiederfindensstrategie zu verwenden.

Die Analysen der letzten Abschnitte zeigen, dass das Lokalisieren eines Tweets in der eigenen *Timeline* oder in der *Timeline* eines Twitter-Nutzers, also dessen Profilseite, am meisten Zustimmung von den Teilnehmern erhalten und somit die beliebtesten Wiederfindensmethoden darstellen. Zudem wurde gezeigt, dass die Wahl der Wiederfindensstrategie lediglich von der Häufigkeit beeinflusst wird, mit der Nutzer auch Tweets aufbewahren. So konnte festgestellt werden, dass Nutzer, die häufig aufbewahren, signifikant stärker zustimmen, die Methoden *Scan Own Timeline*, *Scan Favourites List* und *Scan External Store* zu verwenden, als die Gruppe der Probanden, die seltener oder nie Tweets aufbewahrt. Für alle anderen Bedingungen bzw. Faktoren, die vermeintlich Einfluss auf die Wahl der Wiederfindensstrategie besitzen, konnten keine Unterschiede in den Gruppe ausgemacht werden, respektive ergibt sich hier ein kohärentes Bild: Die Nutzergruppen, die das Wiederfinden als einfach charakterisieren, die die twitterinternen Features für Re-finding als ausreichend bewerten und noch nie ein frustrierendes Wiederfindenserlebnis erfahren haben, stimmen stets signifikant stärker den Aussagen zur Verwendung aller sechs Wiederfindensstrategien zu. Mit anderen Worten: Nutzer, die alle Methoden einsetzen, um Tweets wiederzufinden, charakterisieren Wiederfinden als einfach, bewerten die Twitter-Features als zufriedenstellend und waren noch nie frustriert.

Einige dieser Beobachtungen werden zusätzlich durch starke Korrelationskoeffizienten (Spearman  $\rho$ ) der Nutzeraussagen gestützt, welche in der Korrelationsmatrix in Abbildung 5.23 dargestellt





**Abbildung 5.23.:** Korrelationsmatrix für Häufigkeits- und Bewertungsvariablen zu Aufbewahren und Wiederfinden auf Twitter

sind. Für empirische Fragebogendaten relativ starke Korrelationen ergeben sich bei der Häufigkeit des Aufbewahrens und der Häufigkeit des Wiederfindens ( $\rho = 0,36$ ), speziell dem Wiederfinden von eigenen gesendeten Tweets ( $\rho = 0,39$ ). Dies steht in leichter Diskrepanz zu der Tatsache, dass Nutzer, die häufig aufbewahren, nicht automatisch auch häufig wiederfinden, wie dies weiter oben beschrieben wurde. Grundsätzlich stützt die Korrelation aber die Aussage, dass Nutzer, die häufiger aufbewahren, auch häufiger wiederfinden. Ferner ist zu erkennen, dass Teilnehmer, welche die Aufbewahrungsoptionen von Twitter besser bewerten auch die Wiederfindensoptionen besser bewerten, wodurch sich ein hoher Korrelationskoeffizient von  $\rho = 0,62$  ergibt. Ebenfalls sehr stark korreliert die Zufriedenheit mit den Wiederfindensoptionen und die Einfachheit des Wiederfindens ( $\rho = 0,50$ ), was nur konsequent erscheint. Eine äußerst starke Korrelation ergibt sich ebenfalls bei der Frequenz des Wiederfindens eigener Tweets und der Häufigkeit von frustrierenden Re-finding-Erlebnissen ( $\rho = 0,48$ ). Dies ist in der Tat überraschend, da eigene Tweets vermeintlich leicht über die eigene Nutzerprofilseite wiedergefunden werden können. Ein Erklärungsansatz für diese starke Korrelation könnte eine Missinterpretation der Frage sein, indem der Frageteil [...] *find one of your own Tweets again* falsch interpretiert wurde und die Studienteilnehmer die Tweets der regulären *Timeline* meinten. Ein weiterer Erklärungsansatz könnte darin bestehen, dass es weniger das Wiederfinden an sich ist, das als frustrierend wahrgenommen wird, als die Beschränkung der zusätzlichen Handlungen, die dann ausgeführt werden können. So gibt es seitens der Twitter-Nutzer schon lange das Bedürfnis nach Änderung/Korrektur von bereits gesendeten Tweets, welches aber bis dato noch unerfüllt ist (K. Rogers, 2016). Solange diese Möglichkeit nicht existiert, sehen sich Nutzer dazu veranlasst, Tweets zu löschen, wie in Abschnitt 3.2.2 beschrieben.

Der folgende Abschnitt widmet sich nun genau dieser Frage: Was macht Wiederfinden auf Twitter zu einem frustrierenden Erlebnis, welche Faktoren spielen dabei eine Rolle und welche Wiederfindensstrategien ermöglichen es, das Frustrationsniveau zu senken?

### 5.7.8. Einflussfaktoren auf das Frustrationsniveau bei Wiederfindensaktionen

58,3% der Probanden berichten davon, schon einmal mit einer frustrierenden Wiederfindensaufgabe konfrontiert worden zu sein. Für 64,9% treten frustrierende Situationen aber nur gelegentlich auf. Mit anderen Worten: einige Studienteilnehmer empfinden Wiederfinden als frustrierend, wohingegen andere diese Erfahrung nicht teilen. Somit stellt sich die Frage, welche Faktoren besonders dazu beitragen, dass Re-finding seitens der Nutzer als frustrierend wahrgenommen wird. Um zu eruieren, welche Faktoren die Wahrnehmung der Probanden beeinflussen und ob bestimmte Merkmale als Prädiktoren für ein frustrierendes Wiederfindenserlebnis genutzt werden können, wurden die Fragebogenantworten in einem logistischen Regressionsmodell verwendet. Dabei wurde das Problem als binäres Klassifikationsproblem interpretiert, ob Nutzer jemals eine Wiederfindenssuchaufgabe als frustrierend empfunden haben oder nicht. Im Folgenden wird ausgeführt, weshalb die logistische Regression ein adäquates Modellbildungsverfahren darstellt, welche Faktoren bei der Modellbildung verwendet wurden und wie die Modellbildung konkret umgesetzt wurde. Schließlich wird das Ergebnis der Modellbildung vorgestellt.

#### 5.7.8.1. Logistische Regression als Analysemethode

Nach Gelman und Hill stellt die logistische Regression die Standardmethode zur Modellierung binärer Ereignisse dar (Gelman & Hill, 2007, S.79). Als Datenanalysemethode liegt ihr Hauptziel darin, die Rolle von Variablen bei der Erklärung einer abhängigen Variable zu verstehen (Hastie et al., 2009, S.121). Hair, Black und Babin begründen die Beliebtheit der logistischen Regression in ihrer Voraussetzungsfreiheit in Bezug auf die Verteilung der Daten sowie geringen bzw. nicht vorhandenen Annahmen über die Verteilung der Daten in den Gruppen (Hair et al., 2010, S.413). Zudem wird sie aufgrund ihrer Ähnlichkeit zur klassischen Regression, den ähnlichen statistischen Test, der Handhabung von metrischen und nominal-skalierten Variablen sowie den Möglichkeiten, um die Passung des Modells an die Daten (*goodness of fit*) zu bestimmen, auch in Situationen bevorzugt, in denen die Voraussetzungen der Daten für parametrische Verfahren gegeben wären (Hair et al., 2010, S.413f). Die logistische Regression ist ein Spezialfall der *generalized linear models* und wird vornehmlich bei binären abhängigen Variablen verwendet, wobei für jede Beobachtung auf der Basis einer oder mehrerer Prädiktoren/unabhängiger Variablen eine Wahrscheinlichkeit für die abhängige Variable zwischen 0 und 1 für die Zugehörigkeit zu einer Klasse/Gruppe vorhergesagt wird (Gelman & Hill, 2007, S.109) (Hair et al., 2010, S.414). Für das oben beschriebene Problem, ob Nutzer jemals frustriert gewesen sind, könnte dies — die unabhängige Variable ANZAHL AN FOLLOWER vorausgesetzt — wie folgt formuliert werden:  $Pr(frustrated = yes | FollowingCount)$  abgekürzt als  $p(FollowingCount)$  bzw. allgemeiner als  $p(X)$ . Folglich würde man  $frustrated = yes$  für die Fälle bei denen  $p(FollowingCount) > 0.5$  zutrifft, vorhersagen.

Zur Anpassung des Modells an die Daten wird bei der logistischen Regression keine gerade Funktion verwendet, wie bei der klassischen, linearen Regression, sondern die logistische Funktion, die die kontinuierlichen Werte, die man bei einer linearen Regression vorfinden würde, in Wahrscheinlichkeiten transformiert (Gelman & Hill, 2007, S.80) (James et al., 2013, S.132). Da es natürlicher ist, mit linearen Auswirkungen von Prädiktoren/unabhängigen Variablen in Bezug auf die Wahrscheinlichkeiten zu arbeiten, wird oft die inverse logistische Funktion bzw. Logit-Variante dieser Funktion bevorzugt, welche über Umwandlungen (Darstellung der Wahrscheinlichkeiten als Chancenverhältnis) auf der Basis des Odds-Verhältnis aus der logistischen Funktion erstellt werden

Art des logistischen Koeffizienten	Veränderung auf Skala
Original	Logit (Log des Chancenverhältnisses (Odds))
Exponential	Odds

**Tabelle 5.3.:** Interpretation von logistischen Regressionskoeffizienten bei unterschiedlichen Referenzskalen

kann (James et al., 2013, S.132). Durch Umwandlung von

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

erhält man

$$\frac{p(X)}{1 - p(X)} = e^{\beta_0 + \beta_1 X}$$

Durch weitere Umwandlung erhält man das Log-Odds auf der linken Seite der Gleichung und den linearen Zusammenhang der Variablen rechts.

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X$$

Regressionskoeffizienten werden in der logistischen Regression über die Maximum-likelihood-Methode auf Basis der Beobachtungen geschätzt. Wie bei der linearen Regression werden Koeffizienten als statistisch signifikant interpretiert, wenn dessen Abstand mindestens zwei Standardabweichungen weg von 0 ist. Die Interpretation der Regressionskoeffizienten ist aufgrund der S-förmigen Funktion und des nicht linearen Zusammenhangs zwischen Prädiktoren und abhängiger Variable komplexer als bei der linearen Regression. Zudem können Regressionskoeffizienten in Abhängigkeit von der verwendeten Software als *odds* oder in ihrer Logit-Form angegeben werden. Hair, Black und Babin zeigen in einer Tabelle (vgl. Tabelle 5.3), wie logistische Koeffizienten angegeben sein können und welche Veränderungen diese Angaben widerspiegeln. Besitzt man die Ausgabe eines logistischen Koeffizienten als *original* muss davon der *logit* genommen werden, um diesen Wert bzw. das Ergebnis auf der Wahrscheinlichkeitsskala interpretieren zu können. Die folgenden Ausführungen beziehen sich auf die Angabe des Koeffizienten im Original. Die Richtung eines Koeffizienten bzw. die Beziehung zwischen unabhängiger Variable und abhängiger Variable lässt sich wie bei der klassischen Regression interpretieren. Handelt es sich um einen positiven Koeffizienten, steigt die Wahrscheinlichkeit. Handelt es sich um einen negativen Koeffizienten, verringert dies die Wahrscheinlichkeit (Hair et al., 2010, S.422). Eine einfache Variante, die schnellen Einblick über die Richtung und Stärke des Koeffizienten einer unabhängigen Variable bietet, ist die sogenannte Divide-by-4-Regel deren Bedeutung kurz geklärt wird.

#### 5.7.8.2. Die Divide-by-4-Regel

Zur Bestimmung des marginalen Effekts, also des Effekts, den eine unabhängige Variable bei einer multiplen logistischen Regression auf die abhängige Variable besitzt, kann die Divide-by-4-Regel genutzt werden. Hierbei wird die betreffende Variable um eine Einheit verändert, während die anderen Variablen konstant gehalten werden. Gelman und Hill erläutern die Divide-by-4-Regel wie folgt (Gelman & Hill, 2007, S.82): Die logistische Kurve ist in ihrem Mittelpunkt am steilsten. An diesem Punkt gilt folgender Umstand:  $\alpha + \beta x = 0$  also  $\text{logit}^{-1}(\alpha + \beta)x = 0.5$ . Die Steigung der Kurve (Ableitung der logistischen Funktion) ist an diesem Punkt maximiert und besitzt den Wert  $\frac{\beta e^0}{(1+e^0)^2} = \frac{\beta}{4}$ . Hieraus folgt, dass das obere Limit des marginalen Effektes einer unabhängigen

Variable stets  $\frac{\beta}{4}$  ist. Gelman und Hill formulieren dies wie folgt: „As a rule of convenience, we can take logistic regression coefficients (other than the constant term) and divide them by 4 to get an upper bound of the predictive difference corresponding to a unit difference in  $x$ “ (Gelman & Hill, 2007, S.82). Bei der Darstellung der Ergebnisse werden die Effekte der Koeffizienten mithilfe dieser Methode in Wahrscheinlichkeiten übersetzt.

### 5.7.8.3. Bestimmung der Anpassungsgüte

Ein wesentlicher Schritt bei der Bildung eines logistischen Regressionsmodells ist das *model assessment* also die Bestimmung des *goodness of fit* oder Anpassungsgüte. Die Anpassungsgüte beschreibt, wie gut ein statistisches Modell die Beobachtungen vorhersagen bzw. erklären kann (Kuckartz et al., 2013, S.278). Nach Hair, Black und Babin gibt es zwei Möglichkeiten den *goodness of fit* eines logistischen Regressionsmodells zu bestimmen (Hair et al., 2010, S.419): Pseudo- $R^2$ -Werte sowie die Vorhersagegenauigkeit (*accuracy*). Beide betrachten die Passung eines Modells aus unterschiedlichen Perspektiven, sollten aber zu identischen Schlussfolgerungen führen. Zur Berechnung von Pseudo- $R^2$ -Werten wird der Likelihood-Wert, genauer der „negative zweifache Log-Likelihood-Wert“ (-2LL-Wert) des Null-Modells (*baseline*, ein Modell ohne Prädiktoren), sowie des vorgeschlagenen neuen Modells verrechnet (Kuckartz et al., 2013, S.278). Die -2LL-Werte geben grundsätzlich Auskunft darüber, wie gut die *maximum likelihood estimation* geschätzt hat. Der -2LL-Wert einer Anpassungsgüte geht gegen 0, was bedeutet, dass kleinere Werte auch eine bessere Modellanpassung bedeuten. Wie bei der linearen Regression gibt der Pseudo- $R^2$ -Wert Auskunft über den Anteil der erklärten Varianz in den Daten (Kuckartz et al., 2013, S.287). Nach Kuckartz sind für die logistische Regression vor allem folgende  $R^2$  Werte gebräuchlich: Cox und Snells  $R^2$  und Nagelkerke  $R^2$ . Erreicht ein Modell Pseudo- $R^2$ -Werte ab 0,4, wird dies als gute Modellpassung interpretiert (Kuckartz et al., 2013, S.279).

### 5.7.8.4. Beschreibung der verwendeten Features

Zur Modellbildung wurden insgesamt 41 Merkmale verwendet, die gemäß der fünf Abschnitte des Fragebogens in fünf Gruppen eingeteilt werden können. Auf ein gezieltes *feature engineering* wurde verzichtet. Bei numerischen Werten wurde lediglich versucht, eine Logarithmierung vorzunehmen, um der Tatsache Rechnung zu tragen, dass bestimmte Zusammenhänge nicht linear sein müssen. Tabelle 5.4 listet die verwendeten Variablen auf.

### 5.7.8.5. Datenanalyse und Modellbildung

Welche Faktoren haben Einfluss auf die Frustrationserfahrung beim Wiederfinden von Tweets? Zu Beginn der Datenanalyse wurde ein automatisches Modellbildungsverfahren benutzt, bei dem alle 41 Features mit dem Ziel betrachtet wurden, signifikante Beziehungen zwischen Features und der abhängigen Variable (Frustrationserfahrung) zu bestimmen. Das Verfahren ergänzt bzw. entfernt Faktoren Schritt für Schritt und beurteilt den Einfluss der hinzugefügten/entfernten Faktoren, indem die Anpassungsgüte (*goodness of fit*) des Modells nach jedem Schritt bestimmt wird. Das Prüfen der Modellpassung nach jedem Schritt wird auch *analysis of deviance* genannt (Hastie et al., 2009, S.124). Die mit dieser Methodik identifizierten Faktoren wurden zusätzlich um weitere Faktoren ergänzt, bei denen vermutet wurde, dass sie Re-finding-Erfahrung frustrierender bzw. weniger frustrierend gestalten könnten. In Anlehnung an die Vorgehensweise von Kabacof wurde nach jeder Ergänzung eines Features ein Chi-Quadrat-Test durchgeführt, um zu überprüfen, ob das hinzugefügte Feature das Gesamtmodell verbessert, indem die *residual deviance* signifikant verkleinert wird (Kabacoff, 2010, S.320). Neben einfachen Features wurden auch Interaktionen zwischen

Name	Beschreibung
<i>Twitter-Accountinformationen und Verhalten</i>	
X1demo_firstpost	Zeit seit dem ersten Tweet von Nutzer Y
X1demo_posttweet	Häufigkeit Tweets Senden von Nutzer Y
X1demo_tweetcount*	Anzahl gesendeter Tweets von Nutzer Y
X1demo_followercount	Anzahl an Follower von Nutzer Y
X1demo_followingcount*	Anzahl an Accounts die Nutzer Y folgt
<i>Favorisieren-Verhalten</i>	
X2fav_favcount*	Anzahl an favorisierten Tweets von Nutzer Y
X2fav_favfreq	Häufigkeit Tweets favorisieren
X2fav_favlistfreq	Häufigkeit Favoritenlistenbesuch
X2fav_favfreqown	Häufigkeit eigene Tweets favorisieren
X2fav_unfavfreq	Häufigkeit Tweets Entfavorisieren
<i>Aufbewahrens-Verhalten</i>	
X3preserve_freq	Häufigkeit Tweets aufbewahren
X3preserve_fav	Zustimmung zur Nutzung der FAV-Funktion
X3preserve_retweet	Zustimmung zur Nutzung der RT-Funktion
X3preserve_addstore	Zustimmung zur Nutzung eines externen Speicherorts
X3preserve_notreat	Zustimmung keine besondere Handlung vorzunehmen
X3preserve_freqown	Häufigkeit eigene Tweets aufbewahren
X3preserve_favown	Zustimmung zur Nutzung der FAV-Funktion
X3preserve_retweetown	Zustimmung zur Nutzung der RT-Funktion
X3preserve_storeown	Zustimmung zur Nutzung eines externen Speicherorts
X3preserve_notreatown	Zustimmung keine besondere Handlung vorzunehmen
<i>Wiederfindens-Verhalten</i>	
X4refind_freq	Häufigkeit Tweets wiederfinden
X4refind_searchTimeline	Zustimmung zur Suche in eigenen Timeline
X4refind_searchfavlist	Zustimmung zur Suche in Favoritenliste
X4refind_searchTimeline_person	Zustimmung zur Suche auf Nutzerprofilseite
X4refind_query	Zustimmung zur Nutzung der Twitter-Suche
X4refind_searchengine	Zustimmung zur Nutzung externer Suchmaschinen
X4refind_lookstore	Zustimmung zur Suche in externem Speicherort
X4refind_freqown	Häufigkeit eigene Tweets wiederfinden
X4refind_searchTimeline_own	Zustimmung zur Suche in eigenen Timeline
X4refind_searchfavlist_own	Zustimmung zur Suche in Favoritenliste
X4refind_query_own	Zustimmung zur Nutzung der Twitter-Suche
X4refind_searchengine_own	Zustimmung zur Nutzung externer Suchmaschinen
X4refind_lookstore_own	Zustimmung zur Suche in externem Speicherort
X4refind_reason_content	Wiederfindensgrund: Tweetinhalt
X4refind_reason_link	Wiederfindensgrund: Link
X4refind_reason_othermedia	Wiederfindensgrund: Anhang(Foto/Video)
X4refind_reason_personRL	Wiederfindensgrund: Real life Bekanntschaft
X4refind_difficulty	Zustimmung zur Aussage: Wiederfinden ist einfach
<i>Zusammenfassung</i>	
X5sum_preservingopt	Bewertung der Aufbewahrungsoptionen
X5sum_refindingopt	Bewertung der Wiederfindensoptionen

**Tabelle 5.4.:** Die 41 Merkmale, die zur Vorhersage des Frustrationsniveaus genutzt werden. Features mit \* besitzen einen normalen und logarithmierten Wert

signifikanten Features betrachtet, wie es die Modellbildungsrichtlinien von Gelman und Hill empfehlen (Gelman & Hill, 2007, S.69ff.). Zur Evaluation der Anpassungsgüte des Modells und seiner Vorhersageleistung werden mehrere Varianten von Pseudo- $R^2$ -Maße, der C-Index-Wert (*bias corrected AUC value*), sowie weitere Klassifikationsmaßzahlen wie die Genauigkeit der Vorhersage, als Gütekriterien angeführt.

#### 5.7.8.6. Ergebnis: Logistisches Regressionsmodell

Tabelle 5.5 zeigt das finale Regressionsmodell und offenbart diejenigen Features, die dabei helfen, festzustellen, ob ein Nutzer bereits einmal ein frustrierendes Wiederfindenserlebnis erfahren hat oder nicht. Das Vorzeichen sowie die statistische Signifikanz des INTERCEPT sind nicht von Bedeutung und können in der weiteren Betrachtung vernachlässigt werden (Gelman & Hill, 2007, S.83).

Das Feature FOLLOWINGCOUNT(LOG) zeigt, dass die Anzahl an Accounts denen ein Nutzer folgt, einen signifikanten Prädiktor für die Vorhersage der Frustrationserfahrung darstellt. Dabei handelt es sich nicht um einen linearen Zusammenhang, vielmehr kann durch eine Logarithmierung der Followeranzahl eine bessere Vorhersage erzeugt werden. Der Regressionskoeffizient zeigt ein po-

KOEFFIZIENTEN	Estimate	Std. Error	z value	Pr(>  z )
(Intercept)	-8.56255	0.98035	-8.734	< 2e-16 ***
Followingcount(log)	0.11382	0.05024	2.266	0.02348 *
Frequency: Favouriting own Tweets	0.23472	0.38212	0.614	0.53904
Frequency: Unfavouriting Tweets	0.50839	0.22065	2.304	0.02122 *
Frequency: Preserving Tweets	1.29886	0.22427	5.791	6.98e-09 ***
Frequency: Re-finding Tweets	1.48256	0.29879	4.962	6.98e-07 ***
Frequency: Re-finding own Tweets	1.01546	0.21728	4.674	2.96e-06 ***
Re-find Method: Search Engine	0.18185	0.10328	1.761	0.07829 .
Re-find Method: Search <i>Timeline</i> Person	0.31484	0.11486	2.741	0.00612 **
Interaction:				
Frequency: Favouriting own Tweets & Frequency: Preserving Tweets	-0.23670	0.10618	-2.229	0.02580 *
Interaction:				
Frequency: Preserving Tweets & Frequency: Re-finding	-0.23437	0.10463	2.240	0.02509 *

Significance codes: 0 '\*\*\*'0.001 '\*\*'0.01 '\*'0.05 '.'0.1 |Null deviance: 821.35  
Residual deviance: 600.75 |AIC: 622.75

Tabelle 5.5.: Koeffizienten des logistischen Regressionsmodells

sitives Vorzeichen, wodurch deutlich wird, dass eine höhere Anzahl an Follower in einer höheren Wahrscheinlichkeit resultiert, bereits einmal frustriert gewesen zu sein. Das Feature FREQUENCY: FAVOURITING OWN TWEETS, also die Häufigkeit mit der eigene Tweets favorisiert werden, — welches in vorausgegangenen Abschnitten als eine beliebte Methode zur Aufbewahrung von Tweets eruiert wurde — ist nicht signifikant, besitzt aber das richtige Vorzeichen und ist Teil einer signifikanten Interaktion. Folgt man den Modellbildungsprinzipien nach Gelman und Hill, können solche Features Teil des gesamten Modells bleiben, da sie die Vorhersage nicht verschlechtern (Gelman & Hill, 2007, S.69).

Betrachtet man die Interaktion mit der Variable FREQUENCY: PRESERVING TWEETS so besitzt FREQUENCY: FAVOURITING OWN TWEETS insgesamt gesehen sogar einen negativen Effekt im Modell. Wendet man die Divide-by-4-Regel an, also erhöht den Wert der Variable FREQUENCY: FAVOURITING OWN TWEETS und setzt man die beteiligte Variable FREQUENCY: PRESERVING TWEETS auf ihren Durchschnittswert, zeigt sich, dass die Wahrscheinlichkeit, dass Nutzer bereits einmal durch Wiederfinden frustriert waren, um 8% abnimmt<sup>21</sup>. Das Gegenteil kann für das Entfavorisieren von Tweets beobachtet werden: erhöht man die Frequenz des Entfavorisierens auf der Likert-Skala um einen Wert, so steigt die Wahrscheinlichkeit, dass Nutzer bereits frustriert waren, um 12,5%. Generell zeigt das Modell jedoch, dass je häufiger Nutzer Tweets aufbewahren (FREQUENCY: PRESERVING TWEETS), desto stärker nimmt die Wahrscheinlichkeit zu, dass sie bereits einmal frustriert waren. Dieser Unterschied wird zusätzlich durch die Interaktion der beiden Variablen deutlich, da der Einfluss des Prädiktors FREQUENCY: PRESERVING TWEETS bei denjenigen Probanden an Bedeutung verliert, die häufig eigene Tweets favorisieren. Genau wie häufiges Aufbewahren zu erhöhtem Frustrationsrisiko führt, führt auch häufiges Wiederfinden bei den Nutzern zu einer gesteigerten Wahrscheinlichkeit Frustration erfahren zu haben. Die beiden Variablen FREQUENCY: RE-FINDING TWEETS und FREQUENCY: RE-FINDING OWN TWEETS, welche die Häufigkeit repräsentieren, mit der Nutzer Tweets wiederfinden, besitzen beide positive Koeffizienten, woraus erkenntlich wird, dass je häufiger sich Nutzer mit Re-finding auseinandersetzen, desto wahrscheinlicher ist es, dass sie bereits einmal ein frustrierendes Wiederfindenserlebnis gehabt haben. Der negative Koeffizient der Interaktion zwischen FREQUENCY: PRESERVING TWEETS und FREQUENCY: RE-FINDING TWEETS verdeutlicht aber, dass die Bedeutung des jeweiligen anderen Features geringer wird, falls Nutzer häufiger aufbewahren oder häufiger wiederfinden. Die Bedeutung des Prädiktors FREQUENCY: PRESERVING TWEETS nimmt folglich ab, wenn Nutzer häufiger wiederfinden und vice versa.

<sup>21</sup>Der Durchschnitt von *Frequency: Preserving* = 2.344. Setzt man in die Interaktion die beiden beteiligten Koeffizienten ein, ergibt sich folgende Gleichung:  $0,234 - 0,236 * 2,344 = -0,319$ .  $\frac{-0,319}{4} = 0,079$ .

Das Regressionsmodell beinhaltet zwei Wiederfindensmethoden: Wiederfinden via externer Suchmaschine (RE-FIND METHOD: SEARCH ENGINE) und Wiederfinden via Suche auf der Profilseite/den gesendeten Tweets eines Nutzers (RE-FIND METHOD: SEARCH *Timeline* PERSON). Die Koeffizienten beider Methoden besitzen ein positives Vorzeichen, wodurch sich eine stärkere Zustimmung zur Verwendung dieser Methoden in einer höheren Wahrscheinlichkeit von Frustrationserfahrung niederschlägt. Wendet man die Divide-by-4-Regel an, zeigt sich, dass das Wiederfinden via Suchmaschine in einer 3% höheren Wahrscheinlichkeit und das Wiederfinden via Nutzerprofilseite in einer 7% gesteigerten Wahrscheinlichkeit bereits Frustration erfahren zu haben, resultiert. Dabei sind die Ergebnisse nicht als kausale Zusammenhänge zu interpretieren in dem Sinne, dass Nutzer, die diese Methoden verwenden, frustriert zurückgelassen werden. Vielmehr deutet dieses Ergebnis darauf hin, dass Nutzer, die bereits einmal frustriert worden sind, wenn sie einen Tweet nicht wiederfinden konnten, oder der Prozess unangenehm lange gedauert hat, fortgeschrittene Methoden des Wiederfindens verwenden, wie die Nutzung einer externen Suchmaschine.

PSEUDO $R^2$ MEASURES:		BIAS CORRECTED C-INDEX:	CLASSIFICATION ACCURACY:	
Cox and Snell Index	0.306	0.823	Precision	77.88
Nagelkerke Index	0.411		Recall	75.70
McFadden's $R^2$	0.268		Accuracy	73.76
			F1	76.77

**Tabelle 5.6.:** Überblick über Maße der Anpassungsgüte sowie die Klassifikationsleistung des Modells

Tabelle 5.6 zeigt den *goodness of fit* des Regressionsmodells. Der Pseudo- $R^2$ -Wert von Nagelkerke zeigt, dass das Modell eine gute Anpassungsgüte vorweisen kann. Der *concordance index* (C-Index) ist vergleichbar mit dem AUC-Wert<sup>22</sup> und ist ein nützliches Maß, um die Güte einer Vorhersageleistung zu bestimmen. Ein C-Index-Wert von 1 würde eine optimale Vorhersage bedeuten, wohingegen ein C-Index-Wert von 0,5 einem zufälligen Klassifikator entspricht. Im vorliegenden Fall wurde mit Hilfe des Bootstrapping-Verfahrens der überoptimistische *bias* des Modells verringert, um eine Bias-corrected-Variante des C-Index zu erlangen, welcher mit einem Wert von 0,8 dem Modell eine sehr gute Vorhersageleistung bescheinigt. Eine Beschreibung dieser Bootstrap-Vorgehensweise zur Berechnung des C-Index findet sich bei Frank Harrell (Harrell, 2006). Auch die benannten Maße zur Klassifikationsleistung zeigen ein einheitliches Bild: bei 3 von 4 Probanden ist das Modell dazu in der Lage richtig vorherzusagen, ob sie bereits frustriert worden sind oder nicht. Tatsächlich kann eine höhere Genauigkeit erreicht werden, was bedeutet, dass diejenigen, die als frustriert vorhergesagt werden auch tatsächlich einmal frustriert waren, jedoch nicht alle Fälle mit Frustrationserfahrung erkannt werden (geringerer Recall).

#### 5.7.8.7. Fazit zum Regressionsmodell

Viele der letzten Abschnitte haben gezeigt, dass Nutzer zu bereits gesehenen Tweets zurückkehren und Tweets per se nicht derart kurzlebig sind, wie man das von Social-Media-Beiträgen erwarten könnte. Die Analysen zeigen zudem, dass das Zurückkehren zu Tweets frustrierend sein kann, was aber per se nicht immer der Fall sein muss. Hierbei liegen die Ergebnisse in Einklang mit früheren PIM-Studien aus Kontexten wie Desktop-Suche oder E-Mail-Suche, die gezeigt haben, dass Re-finding-Aufgaben einfach und geradlinig verlaufen können, manchmal aber auch schwierig, komplex und frustrierend sein können (Dumais et al., 2003; Teevan, 2007; Elswiler, Baillie

<sup>22</sup>Area under the receiver operator characteristic curve (ROC).

& Ruthven, 2011). Dies kann im vorliegenden Fall auch für Twitter bestätigt werden. Fasst man die Erkenntnisse des Regressionsmodells zusammenfassen, ergibt sich ein klares Bild: sehr aktive Nutzer, die vielen Accounts folgen, häufig Tweets aufbewahren und wiederfinden (dies auch mit eigenen Tweets durchführen) und alternative Methoden zum Wiederfinden verwenden, besitzen eine höhere Wahrscheinlichkeit, bereits einmal frustriert worden zu sein. Auch Gomez-Rodriguez und Kollegen haben gezeigt, dass Nutzer, die vielen Accounts folgen, sich deutlich anders verhalten als Nutzer mit weniger Followees, was sie durch einen *information overload* der Nutzer begründen (Gomez-Rodriguez et al., 2014). Darüber hinaus zeigt sich im vorliegenden Regressionsmodell — ähnlich wie bei Studien zum Wiederfindensverhalten mit E-Mails bei denen gezeigt wurde, dass Nutzer, die häufig E-Mails in Ordner aufbewahren bzw. kategorisieren diese Aufgabe als schwierig charakterisiert haben — dass Nutzer, die aufbewahren, auch wahrscheinlicher bereits ein frustrierendes Erlebnis hatten (Elsweiler, Baillie & Ruthven, 2011). Die Nutzung der Favoriten-Funktion stellt hier eine Ausnahme dar, da sie — folgt man dem Modell — die einzige Aufbewahrungsmethode ist, die es ermöglicht, die Wahrscheinlichkeit frustriert worden zu sein, zu verringern.

## 5.8. Analyse von Re-finding-Motiven: Warum Wiederfinden?

Die Frage nach den Gründen und Motiven, die Nutzer dazu bewegen, Tweets wiederzufinden, ist eine zentrale Frage für das Wiederfindensverhalten auf Twitter. Aufgrund ihrer Bedeutung wurde versucht, eine größtmögliche Datenmenge an Wiederfindensgründen aufzubauen. Hierfür wurden die qualitativen Daten aus mehreren Studien zu einer großen Datenbasis zusammengeführt. Die Datenbasis besteht aus folgenden Teilen:

- Die Antworten auf die Freitextfragen 4.3 *Can you name reasons that motivate you to find previously seen Tweets again?* und 4.7 *Can you name reasons that motivate you to find your own Tweets again?* der Twitter-Umfrage.
- Die Antworten auf die Fragen 3. bzw. 4. *If it was re-finding, what was your motivation to get back to that Tweet?* aus dem *Experience Sampling* der Twitter-Logstudie (vgl. Abschnitt 6.3.2.2 bzw. Abbildung 6.6).
- Die Gründe für die Nutzung der Twitter-Suche aus einer Studie von Elsweiler und Harvey, bei der Nutzer explizit angaben, dass die Hauptmotivation für die Suche Re-finding war (Elsweiler & Harvey, 2015).

Aus diesen drei Quellen ergibt sich eine Datenbasis aus 372 Wiederfindensgründen, auf welche derselbe Auswertungsprozess angewendet wurde, wie er in Abschnitt 5.6.1 für die Favorisierungsgründe beschrieben ist.<sup>23</sup> In einem ersten Schritt wurde in einem kollaborativen Prozess ein Kodierschema entwickelt. Bei der Entwicklung des Schemas waren drei Personen beteiligt, wobei 50 Gründe zufällig ausgewählt wurden, anhand derer das Schema entwickelt wurde. Anschließend wurde dieses Schema evaluiert, indem erneut 50 Wiederfindensgründe aus der Datenbasis ausgewählt wurden und dieselben drei Personen eine Probe-Kodierung durchgeführt haben. Zur Überprüfung der Reliabilität bzw. der Urteilsübereinstimmung wurde das Cohens Kappa sowie die Intraklassenkorrelation berechnet. Für Cohens Kappa ergeben sich dabei folgende paarweise Werte, welche der Tabelle 5.7 zu entnehmen sind.

Der Mittelwert aus den drei paarweisen Kappa-Werten beträgt  $\kappa = 0,51$ . Der ICC für die drei Kodierenden beträgt ebenfalls:  $ICC = 0,51$ . Nach Landis und Koch kann man in Bezug auf die erste Version des Schemas nur von einer ausreichenden Übereinstimmung zwischen den Kodierenden sprechen (Landis & Koch, 1977). Krippendorff führt sogar eine konservativere Interpretation an

<sup>23</sup>An der Auswertung der Wiederfindensgründe waren Isabella Hastreiter und Thomas Wilhelm mit beteiligt.



	Coder1	Coder2	Coder3
Coder1	1,00	0,54	0,44
Coder2	0,54	1,00	0,55
Coder3	0,44	0,55	1,00

**Tabelle 5.7.:** Paarweise Berechnung des Kappa-Koeffizienten für alle Codes des Kategoriensystems

(Hallgren, 2012, S.29). Demnach sollte bei Werten von  $\kappa < 0,67$  von Schlussfolgerungen in Bezug auf Daten und Schema Abstand genommen werden. Aufgrund dieser Tatsache wurde eine Analyse der vergebenen Codes vorgenommen, um die Diskrepanzen zwischen den Kodierenden aufzudecken. Dabei wurde ersichtlich, dass in 13 von 50 Fällen (26%) lediglich unterschiedliche Sub-Codes vergeben wurden, grundsätzlich aber eine Übereinstimmung in Bezug auf die Oberkategorie auszumachen war. Aus diesem Grund wurden die beiden Koeffizienten zur Beurteilerübereinstimmung  $\kappa$  und  $ICC$  erneut auf der Basis der Code-Oberkategorien berechnet. Tabelle 5.8 zeigt das Ergebnis der paarweisen Berechnungen bei einer Begrenzung des Koderaums auf Code-Oberkategorien.

Es zeigt sich eine deutlich höhere Übereinstimmung, die im Mittelwert für alle drei Kodierenden bei  $\kappa = 0,74$  ( $ICC = 0,74$ ) liegt. In Bezug auf Sub-Codes wurde das Schema dennoch einer leichten Revision unterzogen. Dabei wurden die von Mayring empfohlenen Richtlinien zur Revision von Code-Schemata herangezogen (Mayring, 2014, S.81). Diese lauten:

- Prüfung des Abstraktionsgrades: Wenn zu wenige Kategorien bestehen, kann es sein, dass diese zu allgemein sind. Wenn zu viele Kategorien existieren, kann es sein, dass diese zu speziell sind.
- Werte für Dimensionen/Kategorien werden entfernt, wenn sie zu detailliert sind. Sie werden hinzugefügt, falls ein höherer Detailgrad vorliegen sollte.
- Definitionen von Kategorien werden — falls notwendig — präzisiert.
- Bei Diskrepanzen oder Problemen der Kategorisierung, also in Fällen in denen eine Differenzierung zwischen Codes nicht möglich war, werden passendere Kategorien Codes diskutiert. Auf der Basis dieser Erkenntnisse werden neue Kodierregeln formuliert und in das Kodierschema übertragen.
- Bereits kodierte Elemente können als Referenzbeispiele in das Kategorienschema übernommen werden.

Besonders relevant ist die Richtlinie zum Abstraktionsgrad, der für manche Codes offensichtlich zu niedrig war, so dass eine zu starke Differenzierung der Codes dazu führte, dass kein Konsens zwischen den Kodierenden gefunden werden konnte. Als Konsequenz wurden somit einige Sub-Codes diskutiert und zusammengefasst. Im Anschluss wurde das fertige Kategoriensystem auf die gesamten Daten angewendet. Im Folgenden wird das Kategoriensystem sowie die Verwendungshäufigkeit der einzelnen Codes vorgestellt.

### 5.8.1. Motivation und Beweggründe für das Wiederfinden von Tweets

Tabelle 5.24a zeigt das verwendete Codeschema sowie die detaillierte Aufteilung der einzelnen Codes bzw. Sub-Codes und deren Verwendungshäufigkeit. Insgesamt wurden 522 Codes auf 372 Gründe vergeben, was im Durchschnitt 1,4 Codes pro Wiederfindensgrund entspricht. Drei Wiederfindensgründe erhielten ein Maximum von fünf Codes. Abbildung 5.24b zeigt die Verwendungs-

	Coder1	Coder2	Coder3
Coder1	1,00	0,77	0,70
Coder2	0,77	1,00	0,75
Coder3	0,70	0,75	1,00

**Tabelle 5.8.:** Paarweise Berechnung des Kappa-Koeffizienten auf der Basis der Kode-Oberkategorien

häufigkeit der Codes auf Kodekategorieebene sowie die Codes, die am häufigsten gemeinsam verwendet wurden (*code intersection*). Im Folgenden werden die Kodekategorien bzw. Sub-Kodes der einzelnen Kategorien näher vorgestellt. Wie in Abschnitt 5.6.3 gilt hierbei die Konvention, dass Codes in eckigen Klammern mit der jeweiligen Kodenummer (z. B. vgl. Kode [1.1]) wiedergegeben werden. Wiederfindensgründe sind mit einem „R“ gekennzeichnet und ebenfalls in eckigen Klammern z. B. „wanted to find an old reply to a blog entry of mine“ [R1]<sup>24</sup>.

### [1] Administrative Motives

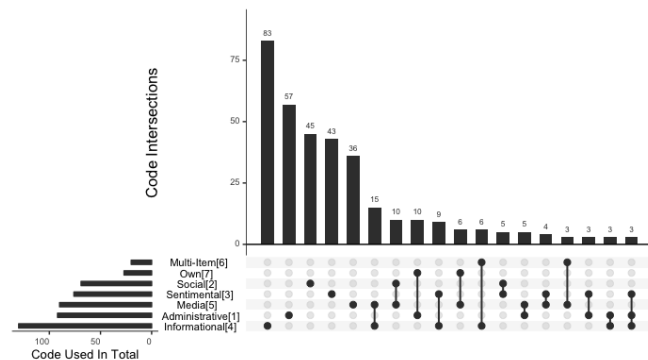
Die Kodekategorie [1] *Administrative Motive* beschreibt Wiederfindensgründe bei denen das Wiederfinden aus einer administrativen, verwaltenden Intention heraus erfolgt. Diese Kodekategorie enthält vier Sub-Kodes und wurde insgesamt 97 mal verwendet. Kode [1.1] *Tweet Composing* findet bei Begründungen Verwendung, die mit dem Verfassen von Tweets in Beziehung stehen. Dies können sowohl eigene verfasste Tweets sein als auch Tweets von anderen Accounts, um auf diese Tweets zu antworten (@Reply). Zu möglichen Gründen gehört die Rückkehr zu Tweets, um sie nach Rechtschreibfehler oder falschen Links zu überprüfen und hierauf entsprechend zu reagieren (z. B. durch das Löschen des Tweets). Kommentare von Probanden wie „[...] look for spelling mistakes“ [R250], „double check the spelling in it“ [R4], „checking for errors, verifying it was really sent“ [R320] oder „To double check that it was correct or accurate“ [R174] belegen dies. Dies beschränkt sich aber nicht rein auf Formalia wie Rechtschreibfehler. Auch das Bedürfnis, die Sinnhaftigkeit eines Tweets zu prüfen, ist bei den Probanden vorhanden: „Check if the tweeted information was stupid or not“ [R243]. Stellen Nutzer bei einem erneuten Prüfen von eigenen gesendeten Tweets fest, die Inhalte sind nicht hochwertig genug oder enthalten heikle, zu persönliche Informationen, so löschen sie diese Nachricht. Dieses Verhalten wird in der Literatur auch als *Regret Deletion* beschrieben und konnte bereits in früheren Arbeiten zu Facebook und Twitter festgestellt werden (Zhao et al., 2013; Sleeper, Cranshaw et al., 2013). Der Wunsch nach einer Editierfunktion, um derartige Tweets nicht mehr löschen zu müssen sondern einfach abändern zu können, ist ein schon lange gehegter Wunsch seitens der Community. Laut einem aktuellen Artikel des *The Telegraph* ist man auch bei Twitter zu der Erkenntnis gekommen, dass diese Funktion notwendig ist, so dass bereits Anstrengungen zu dessen Umsetzung bestehen Titcomb (2016).

Die vorher angeführten Motive sind zeitlich gesehen eher von kurzfristiger Natur und somit möglicherweise als *Re-visits* oder *short term re-finding* zu bezeichnen. Zeitlich etwas versetzt, aber immer noch zu Sub-Kode [1.1] gehörig, ist die Rückkehr zu Tweets, um das erneute Senden eines Inhalts (*duplicate content*) z. B. eines Bildes zu vermeiden „I want to see whether I tweeted a given picture already“ [R86]. Wiederfinden kann aber auch durch das genaue Gegenteil motiviert sein. Nutzer suchen nach einem bereits gesendeten Tweet, um ihn erneut zu versenden, wie dies folgende Kommentare veranschaulichen: „to copy the content because I cannot remember it and lazy to write it again“ [R40], „just want to see them and post them again“ [R264] oder einfach „Recycling“ [R141]. All diese Gründen oder Motive haben zusätzlich eine Gemeinsamkeit: sie beziehen sich fast ausschließlich auf eigene gesendete Tweets. Somit sind alle Begründungen, die hier aufge-

<sup>24</sup>Die Begründungen sind als Zitate zu verstehen, weshalb Fehler jeglicher Art nicht korrigiert wurden

1	<b>Administrative Motives</b>	<b>(5) 97</b>
1.1	Tweet Composing	23
1.2	Monitoring/Checking Status	22
1.3	Specific Use	36
1.4	Time Management	11
2	<b>Social Motives</b>	<b>70</b>
2.1	Share on Platform: RT	9
2.2	Generic Sharing	61
3	<b>Sentimental Motives</b>	<b>(7) 82</b>
3.1	Entertainment	37
3.2	Nostalgia/Memory	22
3.3	Re-experience Other Sentiments	16
4	<b>Informational Motives</b>	<b>(44) 131</b>
4.1	Specific Detail/Item	55
4.2	Specific Topic	32
5	<b>Media Re-finding</b>	<b>94</b>
5.1	Links	53
5.2	Pictures/Videos	24
5.3	Quotes/Comments	17
6	<b>Multi-Item Re-finding</b>	<b>20</b>
7	<b>Own Tweets</b>	<b>27</b>
		$\Sigma = 522$

(a) Kodierschema



(b) Verwendungshäufigkeit von Codes

**Abbildung 5.24.:** Kodierschema (5.24a) der Wiederfindensgründe mit Häufigkeit der Kodennutzung (5.24b)

führt sind, auch mit Kode[7] *Own Tweets* zu versehen. Dies lässt sich auch der Abbildung 5.24b entnehmen. Eine weitere Dimension des Codes [1.1], die nicht mit eigenen Tweets in Beziehung steht, ist die Dimension des Antwortens auf andere Tweets. Dies wurde ebenfalls zu Kode [1.1] hinzugenommen, da die Erstellung eines Tweets im Vordergrund steht. Tweets wiederzufinden, um auf diese zu antworten, ist ein häufig genanntes Motiv und offenbart sich in folgenden Kommentaren: „I want to reply to the person[...]“ [R279], „The one above was because I wanted to reply to the picture [...]“ [R142] oder „[...] reply to a tweet that I hadn’t originally had the time to“ [R161].

Die Begründung [R161] führt eine zusätzlich Dimension der administrativen Kategorie an, nämlich die des Zeitmanagements (*Time Management Relation*), welche durch Kode [1.4] repräsentiert ist. Dieser Kode wurde auf der Basis von Begründungen erstellt, die die Rückkehr zu Tweets mit zu wenig Zeit bei der ersten Begegnung mit selbigem begründen. Beispiele hierfür sind: „I didn’t have the time to click on it“ [R182], „[...]I might also look back on a conversation that had not previously been finished“ [R226] oder „I was looking at the link earlier and decided to visit later when I had more time“ [R23]. Es zeigt sich, dass das Überprüfen der neuesten Tweets in der *Timeline* oft in Situationen geschieht, in denen Nutzer der Social-Media-Applikation nicht ihre völlige Aufmerksamkeit über einen längeren Zeitraum zuwenden können. Treffen sie dann auf besonders relevante Tweets, wird das nähere Begutachten dieser Tweets oft auf einen späteren Zeitpunkt zurückgestellt. Dass hierbei Probleme des Erinnerns dieses Tweets auftreten, liegt in der Natur der Sache und zeigt sich auch in Nutzerkommentaren: „[...] got distracted doing something else and forgot about it until later on“ [R206]. Ein weiterer Sub-Kode aus der Kategorie der administrativen Beweggründe steht in Zusammenhang mit dem Überprüfen des Status oder Erfolges eines eigenen gesendeten Tweets, also ob er häufig weitergeleitet oder favorisiert wurde, oder dem Lesen von Antworten auf diesen Tweet. Diese Motive sind repräsentiert durch Kode [1.2] und spiegeln sich in Kommentaren wieder wie „wanted to see if there were any replies“ [R67], „to see what is responded“ [R220], „looking for responses [...]“ [R250] oder „personal greed“ [R52]. Auch hier werden diese Gründe zumeist mit Kode [7], der für eigene gesendete Tweets steht, begleitet. Der letzte Sub-Kode der Kategorie — Kode [1.3] *Specific (Re)-Use* — beschreibt Wiederfindensgründe, die eine konkrete Handlung mit dem wiedergefundenen Tweet erwähnen, wie folgender Kommentar

veranschaulicht: „I suppose if I wanted to read it later“ [R218]. Diese Handlung kann z. B. auch das spätere Speichern eines Tweets sein: „I remembered that Max tweeted something about a nice building inside the campus. I wanted to save the image.“ [R260] oder „[...]something containing relevant information i wanted to save for later.“ [R50]. Nutzer kehren also zu Tweets zurück, um diese nachträglich noch aufzubewahren. Ein weiteres konkretes Einsatzszenario besteht darin, einen Tweet in einer bestimmten Situation, z. B. in einer Diskussion, zu verwenden: „Needed it in discussion[...]“ [R361], „Use it for later in work / discussions“ [R79]. Ein weiterer Probandenkommentar, der das Wiederfinden eines eigenen Tweets für ein konkretes Einsatzszenario beschreibt, ist folgender: „I wanted to search my own tweet. I know I mentioned which airport I should use to get to Brixen, but forgot. I needed to book my flights.“ [R134].

## **[2] Social Motives: Sharing**

Mit 70 Begründungen, die einen Kode dieser Kategorie besitzen, ist das Weiterleiten oder Teilen ein zentraler Wiederfindensgrund. Kode [1.3] *Specific Use* beschreibt die Nutzung eines Tweets in einer neuen Situation, jedoch liegt der Fokus der Nutzung auf der Seite des Nutzers. Kode-Kategorie [2] *Social Motives* beschreibt eine spezielle Nutzung, die auf andere Personen abzielt, also sozial motiviert ist, nämlich das Teilen eines Tweets mit anderen Personen. Dies kann via RT, also weiterleiten auf Twitter (Kode [2.1] *Share on Platform:RT*), oder auf generische Weise z. B. via E-Mail, anderen Social-Media-Applikation oder realen Situationen erfolgen, wofür Kode [2.2] *Generic Sharing* vorgesehen ist. Nutzerkommentare, die das Wiederfinden mit einem anschließenden RT begründen sind beispielsweise folgende: „i might want to re-tweet them“ [R210], „looking for a link to retweet[...]“ [R139] „I wanted to refind the tweet containing a ted-talk I started watching earlier in the afternoon. The aim being to favourite and then retweet it“ [R251]. Wiederfindensbegründungen, die ein Weiterleiten oder Verbreiten auf generische Art und Weise formulieren bzw. keine konkrete Methodik nennen, gibt es zahlreiche ( $n = 61$ ). Oft werden dabei Verben wie *share* oder *show* in Verbindung mit *friends*, *someone* oder *others* verwendet („want to share them with someone“ [R166], „so i can show them to someone“ [R188], „if it’s something i want to share with someone else not on twitter“ [R104]). In wenigen Fällen werden aber konkrete Off-Plattform-Methoden genannt, z. B. andere Social-Media-Plattformen, wie diesem Probandenkommentar zu entnehmen ist: „[...]Or I forgot to send the tweet to Facebook (via #fb) and want to copy it and post it there.“ [R352]. Die Verbreitung einer Social-Media-Nachricht über mehrere Plattformen hinweg stellt folglich einen Wiederfindensgrund dar, der abermals die Cross-Plattform-Interaktionen bzw. Social-Media-Ökologie der Nutzer betrifft und weitere Untersuchungen bedürfte.

## **[3] Sentimental Motives**

Bei Begründungen der Kode-Kategorie [3] *Sentimental Motives* stehen emotionale Motive bzw. Gefühle im Vordergrund. In insgesamt 82 Kommentaren führen Probanden emotionale Motive als Wiederfindensgrund an. Kode [3.1] *Entertainment* repräsentiert einen sehr häufigen Wiederfindensgrund, nämlich die Tatsache, dass ein Tweet wiedergefunden wird, weil er lustig oder komisch ist, also dem Nutzer zur Unterhaltung dient. Nutzerkommentare sind sehr knapp formuliert wie „because it’s funny“ [R233], „if they’re funny“ oder „i remember a joke i found funny“ [R323]. Oft wird hierbei auch inhaltlich konkretisiert, welche Information oder welches Medium lustig ist, wodurch die Verwendung weiterer Kodes notwendig wird. Dies zeigt sich an Begründungen wie „was looking for a tweet I had seen joking about former yahoo ceo scott thompson’s fake computer science degree“ [R132] oder „Was talking to someone about ADHD, and wanted to share with them a funny tweet id seen once before. had to search google’s index of twitter to find it.“ [R347]. Eine weitere emotionale Dimension in der Kodekategorie entsteht durch Begründungen, die Wiederfinden als nostalgischen Moment charakterisieren (vgl. Kode [3.2] *Nostalgia/Memory*). Zum einen gehören hierzu Kommentare, die das bloße Erinnern als Wiederfindensgrund angeben, wie in Form von „[...] bring back memories [...]“ [R165] oder „something worth remembering“ [R49]. Das Erinnern an sich, das sich Zurückversetzen zu einer bestimmten Situation, wird hier

zum Wiederfindensgrund. Häufig steht dies in Bezug zu eigenen gesendeten Tweets. Im Sinne einer Zeitreise werden eigene ältere Tweets betrachtet, um ein Bedürfnis nach Nostalgie zu befriedigen. Dies zeigen folgende Kommentare sehr deutlich: „[...]to remember what someone else did and when“ [R217], „For general interest (eg. I wonder what I said in that argument, I wonder what my first tweet was, I wonder what photos I have posted etc.)“ [R36] oder „Other times it's just strolling back through my time line to find things I'd retweeted just to enjoy them again (usually not a specific tweet in mind)“ [R54]. Wiederfinden erfolgt in diesem Kontext somit weniger aufgrund eines klar definierten Tweets, der eine bestimmte Information enthält. Vielmehr ist Wiederfinden wie ein Spaziergang durch die eigenen gesendeten Tweets zu interpretieren (*Own Timeline Strolling*), motiviert durch die emotionale Erfahrung, die entsteht, wenn der allererste eigene Tweet gelesen wird, oder man sieht, welche Tweets bzw. Bilder man weitergeleitet oder gesendet hat. Auch eine psychologische Dimension spielt hier eine Rolle. Die Selbstreflexion, also das Nachdenken über sich selbst und wie man sich entwickelt hat, ist in Nutzerkommentaren zu finden: „i like to see things that i have posted in the past to see how much i have changed.“ [R201]. Dass die Selbstreflexion bzw. das Management des Selbstbildes ein Bedürfnis von Social-Media-Nutzer ist, wurde bereits in früheren Arbeiten thematisiert (Hogan, 2010; Zhao et al., 2013). Dass Nutzer hierfür zu den eigenen gesendeten Tweets zurückkehren müssen, liegt auf der Hand.

Es wird deutlich, das Wiederfinden auf Twitter nicht nur ein singuläres Ziel haben muss, sondern auch mehrere Tweets wiedergefunden werden, also *Multi-item Re-finding* betrieben wird. Zudem sind dies oft Tweets, die von einem selber gesendet wurden. Somit sind bei den Begründungen [R36], [R54] und [R201] auch die Codes der Kategorie [6] und [7] notwendig, um das Spektrum der Wiederfindensmotive in diesen Kommentaren zu fassen.

Als letzter Kode dieser Kategorie ist ebenfalls ein eher generischer Kode anzuführen, der Wiederfindensgründe abdeckt, bei dem das Wiederfinden den Zweck verfolgt, eine emotionale Situation erneut zu erfahren. Nutzerkommentare sind beispielsweise: „to inspire me“ [R237], „uplifting or good message“ [R238] oder „it's a sentiment i wanted to see again for inspiration when things were getting tough“ [R341]. Bei diesen Gründen handelt es sich um das erneute Erfahren von bestimmten emotionalen Zuständen, die das Ziel haben, einen zu motivieren oder aufzuheitern.

#### [4] Informational Motives

Mit einer Verwendungshäufigkeit von 131 sind Codes aus Kodekategorie 4 die am häufigsten verwendeten Codes. Abbildung 5.24b zeigt zudem, dass Codes dieser Kategorie, diejenigen Codes sind, die am häufigsten alleine verwendet werden. Enthält die Begründung für das Wiederfinden eine konkrete Information, die der Twitter-Nutzer zu finden intendiert, wie eine Nummer, einen Namen, einen Hashtag oder ähnliches, spricht dies für die Verwendung von Kode [4.1] *Specific Detail/Item*. Enthält der Kommentar eine thematische Ausrichtung bzw. einen Themenbezug wie Sport, Musik, Ernährung etc. dann trifft Kode [4.2] *Specific Topic* zu. Beispiele für konkrete Informationseinheiten zeigen sich in Kommentaren wie „looking for the name of the guy who does recsys user studies“ [R308], „friends bday/anniversary“ [R324], „my husband and i like barter kings and i was telling him that there's a new season and he asked where i heard that.“ [R329] oder „the hashtag i wanted to use in a tweet i was about to send“ [R167]. Bei einigen wenigen Begründungen geben Nutzer an, nach Profilen zu suchen. Aus den Begründungen ist oft nicht erkenntlich, ob es sich um einen Profilnamen handelt, oder ob dieser für die Suche nach einem gesendeten Tweet dieses Nutzers genutzt wird: „Looking to find Tetsuya's twitter profile - didn't know how to spell his surname“ [R223]. In diesen Fällen wurde auch Kode [4.1] vergeben. Begründungen, die keine konkrete Information nennen, aber einen Tweet zu einem spezifischen Themengebiet wiederfinden wollen, sind folgende: „A reference to a business oriented report, that I remembered a specific comment of (specific in my twitter feed but not on web)“ [R99], „good workout routine or recipe“ [R187] oder „i would like to make these cookies.“ [R185]. Der letzte Kommentar deutet

auf das Wiederfinden eines Rezeptes hin, was in der Tat ein häufig auftretender Wiederfindensgrund war. In der Phase des Probekodierens stellte sich heraus, dass es nicht immer eindeutig ist, ob eine konkrete Information (Kode [4.1]) oder eher ein allgemeines Themengebiet (Kode [4.2]) gemeint war, was häufig zu Nichtübereinstimmung zwischen den Kodierenden führte. Dies scheint jedoch in diesem Kontext weniger entscheidend, als die Tatsache, dass es sich um Informationen mit einem bestimmten Themenbezug handelt, die wiedergefunden werden. Häufig werden von den Probanden auch Kommentare gegeben, in denen generisch von *information* gesprochen wird, oder einfach eine Charakterisierung der Information durch Adjektive wie *relevant* oder *important* vorgenommen wird. Immer dann wurde Kode [4] als Oberkategorie verwendet, wie in folgenden Beispielen: „they are just important“ [R328], „if there was information that was important for me“ [R29] oder „Because they contain some relevant information that I don’t remember; it is very much like knowing that an information is written down in a specific book/chapter, and you can look it up whenever you need to“ [R202]. In der letzten Begründung wird Twitter interessanterweise als beständiges Medium charakterisiert, welches immer wieder zum Nachschlagen verwendet werden kann.

#### [5] Media Re-finding

Kodes der Kategorie [5] *Media Re-finding* werden dann verwendet, wenn die Begründung den Tweetinhalt bzw. die Information wie in Kodekategorie [4] in Bezug auf das Medium näher beschreibt ergo: Links [5.1], Bilder und Videos [5.2] oder Zitate bzw. Kommentare [5.3]. Dies ist bei 94 Wiederfindensgründen der Fall. Die Sub-Kodes der Kodekategorie sind sehr einfach anzuwenden, da ihre Passung für eine Begründung oft leicht zu erkennen ist. Kode [5.1] *Links* ist immer dann passend, wenn die Begründung konkret *Links* oder *references* nennt. Dies ist bei folgenden Kommentaren der Fall: „Find a contained link[...]“ [R4], „if I want to visit a link that was in a Tweet“ [R58], „I need to find a link again[...]“ [R98]. Häufig werden auch die hinter den Links stehenden Inhalte präzisiert, wodurch dann auch die Verwendung eines Kodes aus Kategorie [4] obligatorisch wird, wie in den folgenden Begründungen ersichtlich wird: „There was a cool link to Soundcloud.“ [R6] oder „Looking for the paper posted by Michael Bernstein regarding email prediction“ [R62]. Beide Male kann dies als eine konkrete Information angesehen werden, wodurch zusätzlich die Verwendung von Kode [4.1] nötig ist. Analog verhält sich dies zu Kode [5.2], bei dem ein konkretes Medium, wie ein Bild oder Video, in der Begründung erwähnt wird: „[...] if there’s an interesting picture or link that I want to investigate further“ [R236] oder „Trying to find a link to a video that I remember one of my friends had posted. Pretty sure I saw it on Twitter.“ [R128]. Auch hier werden oft mehrere Kodes benötigt. Beispiele hierfür sind: „Pictures I want to show to others“ [R70] bei der sowohl Kode [5.2] für den Inhalt des Tweets als auch Kode [2.2] für das Teilen mit anderen Personen angebracht wird. Begründungen, die explizit Witze, Zitate oder Kommentare nennen, sind folgende: „because it was a profound phrase“ [R247], „it was a catchy saying“ [R93] oder „want to find interesting quote again“ [R284]. Tatsächlich gibt es auch Motive, in denen Nutzer nicht nur einzelne Kommentare wiederfinden wollten, sondern ganze Konversationen nachvollziehen wollten. Hier wurde zusätzlich zu Kode [5.3] auch Kode [6] *Multi-Item Re-finding* vergeben, der das Wiederfinden mehrerer Tweets beschreibt, wie dies bereits auch bei der Beschreibung der Kategorie [3] angesprochen wurde. Nutzer begründen das Wiederfinden von Konversationsverläufen mit Kommentaren wie „To recall a conversation“ [R292], „i want to remind myself of something or find the conversation again and read it“ [R199] oder „[...]i might also look back on a conversation that had not previously been finished.“ [R226].

#### General Interest - [6] Multi-item Re-finding und [7] Own Tweets

Die beiden Kodes der Kategorie [6] *Multi-item Re-finding* und [7] *Own Tweets* sind mehr als Hilfskodes zu verstehen, als dass sie tatsächlich eigenständige Begründungen darstellen. Nichtsdestotrotz sind sie sehr wichtig, um das Spektrum an Wiederfindensgründen und -motiven wiedergeben zu können. Durch Kode [6] — welcher bei 20 Begründungen zutreffend ist — wird deutlich, dass nicht

nur einzelne Tweets wiedergefunden werden, sondern, dass bestimmte Motive das Wiederfinden von mehreren Tweets umfasst. Dies trifft insbesondere bei Kode [5.3] zu, wenn nicht nur einzelne Zitate oder Kommentare gesucht werden, sondern ganze Konversationsverläufe erneut nachvollzogen werden. Noch viel stärker wird die Perspektive *Multi-item Re-finding* bei Begründungen erkennbar, die Kode [3.2] tragen, wenn Nutzer aus emotionalen oder nostalgischen Beweggründen heraus mehrere eigene gesendete Tweets erneut rezipieren möchten. Neben der emotionalen Ebene, abgedeckt durch Kode [3.2] *Nostalgia/Memory*, treten eigene gesendete Tweets in Zusammenhang mit den Kodes aus Kategorie [1] *Administrative Motives* auf. Eigene Tweets werden erneut aufgesucht, um diese auf Fehler zu überprüfen, oder um den Inhalt zu verifizieren und Duplikate zu vermeiden (vgl. Kode [1.1]). Nutzer kehren ebenfalls zu Tweets zurück, um Antworten zu lesen oder den Erfolg des Tweets — durch Nachsehen wie häufig er favorisiert oder weitergeleitet wurde — zu prüfen. Die Suche nach eigenen gesendeten Tweets (Kode [7]) ist durchaus ein häufiges Motiv. Der Kode wird 27 Mal verwendet. Dass diese Dimension der Gründe überhaupt auftritt, ist der Frage 4.7 im Fragebogen zu verdanken, die die Probanden explizit nach Wiederfindensgründe für eigene Tweets befragte. Es ist sehr unwahrscheinlich, dass diese Perspektive des Wiederfindens evident geworden wäre, falls lediglich eine generische Frage zu Wiederfindensgründe, ohne Konkretisierung, gestellt worden wäre.

### 5.8.2. Diskussion der Wiederfindensgründe und Fazit

Das Kode-Schema macht deutlich, dass ein breites Spektrum an unterschiedlichen Beweggründen existiert, die das Wiederfinden bzw. die Rückkehr zu einem bereits gelesenen Tweet motivieren. Nutzer suchen Tweets erneut auf, um einen Link nochmals zu öffnen, über ein Bild zu lachen, oder diesen Tweet zu einem späteren Zeitpunkt anderen Personen zu zeigen. In der Mehrzahl der Fälle ist das Wiederfindensziel ein Tweet mit einer konkreten Information oder ein Tweet zu einem konkreten Thema (zu erkennen an der hohen Zahl an Kodes der Kategorie [4] vgl. Abbildung 5.24b). Dies ist nicht weiter überraschend. Erstaunlich ist jedoch, dass dies nicht obligatorisch ist. Insbesondere Kode [3.2] und das Verhalten, welches oben als „Spaziergang durch die eigene *Timeline*“ beschrieben wurde, hat gezeigt, dass Wiederfinden auch ohne konkretes Ziel stattfindet. Es liegt ein (Informations-) Bedürfnis vor, dass nicht durch einen konkreten Tweet befriedigt werden kann. Die Tatsache, dass kein konkretes Ziel vorliegt bzw. dass kein konkreter Tweet wiedergefunden werden soll sondern während des Wiederfindensprozesses mehrere (eigene gesendete) Tweets betrachtet werden, macht dieses Verhalten zu einem besonderen Aspekt des Wiederfindensverhaltens. Man kann hier von *Multi-item Re-finding* sprechen, einem Informationsbedarf, der auch beim Wiederfinden von E-Mails beobachtet wurde (Elsweiler, Baillie & Ruthven, 2011). *Multi-item Re-finding* tritt auch dann auf, wenn ganze Konversationsverläufe erneut gelesen werden.

Die Häufigkeit, mit der nach eigenen gesendeten Tweets gesucht wird, ist durchaus überraschend. Das Management oder die Rückkehr zu eigenen erstellten Nachrichten ist ein Verhalten, das auch bei Studien zur Facebook-Nutzung evident wurde. Dort werden Nutzer vorgestellt, die ihre eigene *Timeline* als persönliches Archiv wahrnehmen und auch entsprechend nutzen (Zhao et al., 2013; Zhao & Lindley, 2014) (vgl. Abschnitt 3.2.2.1). Auch auf Twitter werden die eigenen gesendeten Tweets in Form einer Selbstreflexion bzw. dem Aufrechterhalten eines Selbstbildes (vgl. Kodegruppe [1]) und als Schwelgen in alten Erinnerungen (vgl. Kode [3.2]) erneut betrachtet. In Hinblick auf die Tatsache, dass Nutzer sehr oft das Bedürfnis haben, eigene gesendete Tweets zu durchstöbern, wäre es nur logisch, Mechanismen zu gestalten, die dieses Bedürfnis besser unterstützen. Der Status quo sieht hier nur das mühevollen Down-Scrollen in der eigenen *Timeline* als Lösung vor (Sleeper, Cranshaw et al., 2013). Eine einfache Lösung könnte die Suche (per Query) in den eigenen Tweets sein. Dieses Feature der Erweiterten-Suche wird jedoch selten genutzt.

Spannend ist dabei auch die zeitliche Dimension des Wiederfindens. Klar ist, dass das Überprüfen der eigenen gesendeten Tweets auf Rechtschreibfehler oder falschen Links zur Kategorie Kurzzeit-Wiederfinden (*Re-visit*) zuzuordnen ist. Ob es sich hier tatsächlich um Wiederfinden handelt, kann diskutiert werden. Eigene gesendete Tweets werden aber auch nach längerer Zeit überprüft, um z. B. das erneute Senden eines Bildes zu vermeiden oder eine *regret deletion* vornehmen, also wenn Nutzer das Senden eines Tweets bedauern und diesen nach einiger Zeit löschen (vgl. Abschnitt 3.2.2). Dass Kurzzeit-Wiederfinden aber auch in Bezug auf Tweets aus der *Timeline* zutreffen kann, zeigt folgender Probandenkommentar: „Looking for a tweet that i just read 30 secs ago. Why is this so hard? I have lots of tweets and it just seems to be lost“ [R11]. Obwohl der gesuchte Tweet vom Probanden erst vor sehr kurzer Zeit gelesen wurde, war es ihm nicht möglich, diesen erneut aufzufinden. Der Kommentar eröffnet noch weitere Perspektiven. Zum einen, dass der Nutzer offenbar an einem gewissen *information overload* leidet, da er sehr viele Tweets in der *Timeline* hat, zum anderen, dass er Wiederfinden auf Twitter als sehr schwierige Aufgabe charakterisiert. Nicht selten scheinen Nutzer beim Überprüfen der neuesten Tweets überwältigt zu sein, so dass es ihnen schwerfällt, genügend Zeit für die relevanten Tweets aufzubringen bzw. die tatsächlich für sie relevanten Tweets überhaupt zu identifizieren. Nicht ohne Grund geben Nutzer an, dass sie beim ersten Aufeinandertreffen mit dem Tweet zu wenig Zeit hatten. Zeitmangel ist somit ein weiterer Grund, der Wiederfinden motiviert und durch Kode [1.4] abgedeckt ist. Die Rückkehr zu Tweets kann dann durch einen spontanen Anstoß (*trigger*) von außen ausgelöst werden, der Nutzer an den jeweiligen Tweet erinnert: „looking for a link I had previously seen in my feed. Abt US politics, don't recall exact issue. Checked bc something else reminded me that I had seen it and became curious.“ [R310]. Manche Begründungen enthalten auch explizit Hinweise zu Wiederfindensstrategien, also Verhaltensweisen, die durch die beschriebenen Motive angestoßen werden. Einige Beispiele sind folgende: „Was talking to someone about ADHD, and wanted to share with them a funny tweet id seen once before. had to search google's index of twitter to find it.“ [R347] oder „I was looking for a friend's profile to find a link he had posted earlier“ [R125]. Zwei Wiederfindensstrategien werden hier genannt, nämlich die Suche nach einem Tweet mithilfe einer herkömmlichen Suchmaschine und die Suche via des Nutzerprofils und den gesendeten Tweets eines Nutzers. Vor allem Letzteres hat sich bereits in Abschnitt 5.7.3 als häufig genutzte Wiederfindensstrategie herauskristallisiert und wird sich ebenfalls in Abschnitt 6.4.4.5 noch als bedeutende Wiederfindensstrategie herausstellen. Um das Wiederfinden von Tweets zu erleichtern, geben Nutzer auch an, diese nach dem Wiederfinden bewusst aufzubewahren (vgl. Kode [1.3]). Für manche Nutzer scheint diese Notwendigkeit aber gar nicht zu bestehen. Sie nehmen Twitter per se schon als Informationsspeicher wahr, an den sie sich wenden können, falls sie eine bestimmte Information suchen: „[...] it is very much like knowing that an information is written down in a specific book/chapter, and you can look it up whenever you need to“ [R202]. Dass sich das Wiederfinden hier durchaus als schwierige Aufgabe gestalten kann, scheint dem Nutzer nicht bewusst zu sein.

Die Analyse der Wiederfindensgründe konnte ein breites Spektrum an Motiven offenbaren, wobei das Wiederfinden von mehreren Tweets im Sinne einer *Twitter Nostalgia* sicherlich als der unerwartetste Grund angesehen werden kann. Auf einer Meta-Ebene liefern die Nutzerkommentare zusätzliche Information wie sie Wiederfinden auf Twitter charakterisieren, wie z. B. der Überfluss an Tweets, die oft zu geringe Zeit, um Tweets als relevant einzuschätzen, bewusst genannte Wiederfindensstrategien oder die Charakterisierung von Twitter als Informationsspeicher.



## 5.9. Fazit zur Fragebogenstudie

Bei der Formulierung der Leitfragen wurde die Erlangung eines umfassenden Überblicks über PIM-Verhalten auf Twitter als zentraler Untersuchungspunkt der Twitter-Umfrage angeführt. Betrachtet man die Erkenntnisse der letzten Abschnitte, wird deutlich, dass die Twitter-Umfrage wesentlich zu einem besseren Verständnis des PIM-Verhaltens auf Twitter beigetragen hat. Möglich wurde dies insbesondere durch die Rekrutierung mehrerer Probanden via Crowdsourcing, die die Voraussetzungen für eine heterogene Stichprobe an Twitter-Nutzern schaffte und somit die Twitter-Umfrage dazu befähigte, Aussagen über das Verhalten von einer breiten Masse an unterschiedlichen Nutzertypen einzuholen. Die Umfrage offenbart folgende zentrale Erkenntnisse:

- Über drei Viertel (77,3%) der Probanden hatten schon einmal das Bedürfnis, Tweets aufzubewahren. Der Anteil der Probanden, der bereits einmal einen Tweet wiederfinden wollte oder wiedergefunden hat, ist mit 72,4% fast ebenso hoch. Positive Korrelationskoeffizienten belegen diesen Trend und offenbaren, dass Probanden, die häufig Tweets aufbewahren, auch häufig Tweets wiederfinden und dies auch bei eigenen Tweets tun. Diese Zahlen zeigen: PIM-Praktiken wie Aufbewahren und Wiederfinden gehören zur Twitter-Nutzung dazu. Twitter ist nicht nur das Echtzeit-Informationsnetzwerk, für das es viele halten, sondern besitzt auch archivische Charakteristika, was sowohl bei den Favorisierungsgründen als auch bei den Wiederfindensmotiven evident wird.
- Zum Wiederfinden verwenden Nutzer vor allem zwei Strategien: Die Strategie *Scan Own Timeline*, also das Lokalisieren des gesuchten Tweets in der eigene *Timeline* z. B. durch Down-Scrolling und die Strategie *Scan Sender Profile*, also das Lokalisieren des gesuchten Tweets via Nutzerprofilseite und dessen gesendeten Tweets. Zudem wird ersichtlich, dass Probanden, die angeben, häufig Tweets aufzubewahren, auch die entsprechenden Speicher, respektive die damit verknüpften Wiederfindensstrategien (*Scan Favourites List* und *Scan External Store*) nutzen. Suchstrategien wie die Verwendung der twitterinternen Suche oder externen Suchmaschinen, wie Google, finden nur geringe Zustimmung unter den Probanden. Schließlich kann festgestellt werden, dass Wiederfinden dann als einfach charakterisiert wird, bzw. die Wiederfindensoptionen der Plattform als ausreichend bewertet werden, wenn Probanden der Verwendung aller angegebenen Wiederfindensstrategien zustimmen.
- Die Favoriten-Funktion ist für das Aufbewahrensverhalten von großer Wichtigkeit. Die Nutzung des Favoriten-Feature als Lesezeichen-Funktion stellt das zweithäufigste Anwendungsmotiv der Funktion dar. Sie wird sogar dazu genutzt, um Sammlungen beliebter Tweets zu erstellen. Man kann jedoch auch feststellen, dass es zahlreiche weitere — auch konträre — Beweggründe gibt, die Favoriten-Funktion zu nutzen. Nutzer besitzen ihre eigenen mentalen Modelle der Favoriten-Funktion und haben somit unterschiedliche Vorstellungen davon, welche Zwecke sie erfüllt. Insgesamt konnten bis zu 25 unterschiedliche Motive bzw. Anwendungszwecke ausgemacht werden.
- Re-finding kann zu einem frustrierenden Sucherlebnis werden, was bereits in anderen Kontexten gezeigt werden konnte und sich auch beim Wiederfinden von Tweets als evident erweist. Mithilfe eines logistischen Regressionsmodells kann gezeigt werden, dass vor allem Nutzer Frustration erfahren, die auf der Plattform sehr aktiv sind, häufig Tweets aufbewahren, wiederfinden und einer hohen Anzahl an Nutzern folgen. Jedoch kann Aufbewahren und insbesondere das Favorisieren eigener Tweets das Frustrationsniveau geringer halten.
- Eine Analyse von Wiederfindensgründen offenbart 19 unterschiedliche Motive, die Wiederfindensverhalten auslösen können. Die Motive reichen von administrativen und eher auf kürzere Perioden ausgerichtete Wiederfindensaktionen, wie dem Überprüfen von gesende-

ten Tweets auf Rechtschreibfehler, bis hin zu einem Spaziergang durch die *Timeline*, bei der mehrere Tweets auf einmal wiedergefunden werden, um ein Bedürfnis nach Nostalgie zu befriedigen.

- Aus methodischer Perspektive hat sich das Werkzeugkastenmodell der qualitativen Inhaltsanalyse nach Margrit Schreier als geeignete Methode bzw. Herangehensweise zur Analyse qualitativer Fragebogendaten erwiesen, die es erlaubt, sowohl Aspekte der klassischen Inhaltsanalyse als auch der *Grounded Theory* im Analyseprozess miteinander zu kombinieren.

Der Abbildung 4.1 aus Kapitel 4, welche das Forschungsdesign visualisiert, ist zu entnehmen, dass die Erkenntnisse bzw. Ergebnispräsentation auch die Konzeptualisierungsphase bzw. Leitfragen der Folgestudie, der Twitter-Logstudie, beeinflussen. Diese Studie bildet den zweiten großen Beitrag der Arbeit und wird im Folgenden dargelegt.

# 6

## Studie II: Twitter-Logstudie

Kapitel 6 charakterisiert Wiederfindensverhalten auf Twitter in Form einer lebensnahen Logdatenerhebung und -auswertung. Die Logdatenerhebung wurde im Zeitraum von Ende März 2015 bis Ende August 2015 durchgeführt.

Zunächst werden Leitfragen für die Twitter-Logstudie formuliert, sowie die verwendete Methodik der Logdatenerhebung und das genutzte Erhebungsinstrument — eine Erweiterung für den Google Chrome Webbrowser, welche eigenhändig zu diesem Zweck implementiert wurde<sup>1</sup> — vorgestellt. Die Ergebnispräsentation beginnt mit der Charakterisierung der Studienteilnehmer, deren allgemeines Twitter-Verhalten, sowie einer Diskussion der Bedeutung des allgemeinen Verhaltens für die Analyse des Wiederfindensverhaltens. Anschließend wird Re-finding-Verhalten auf Twitter detailliert unter unterschiedlichen Gesichtspunkten empirisch beleuchtet. Zunächst wird diskutiert, wie Re-finding in Clickstream-Daten erkannt werden kann und welche Konsequenzen diese Erkennungsvarianten für die darauf aufbauenden Analysen besitzen. Durch Orientierung an den formulierten Leitfragen wird Wiederfinden anhand deskriptiver und inferenzstatistischer Analysen charakterisiert. In einem experimentellen Teil wird gezeigt, dass Wiederfindensverhalten charakteristisch ist und wiederfindenswerte Tweets Eigenschaften besitzen, die zur automatischen Identifikation von Wiederfindenszielen genutzt werden können. Eine Interpretation und Einordnung der in diesem Kapitel entstehenden Erkenntnisse erfolgt schließlich unter Berücksichtigung der bereits bestehenden Forschung zum Wiederfindensverhalten in anderen Domänen und den Erkenntnissen der Umfrage-Studie aus Kapitel 5 in Kapitel 7.

Im Sinne einer publikationsgestützten Dissertation wurden die Inhalte des Abschnitts 6.4.4 bereits einem Bewertungsverfahren (Doppelblindgutachten) unterzogen, in Form einer Veröffentlichung auf der Fachtagung der *39th International ACM SIGIR Conference on Research and Development in Information Retrieval* vorgestellt und sind dort im Tagungsband zur Konferenz in der *ACM Digital Library* veröffentlicht (Meier & Elsweiler, 2016).

---

<sup>1</sup>Probleme, die bei der Implementierung entstanden, wurden durch die Unterstützung von Alexander Bazo und Thomas Wilhelm gelöst.

## 6.1. Leitfragen der Twitter-Logstudie

Aufbauend auf den Resultaten der Twitter-Umfrage können Leitfragen formuliert werden, welche durch eine Clickstream-Datenanalyse unterstützt werden sollen. Getreu dem Prinzip des in Abschnitt 4 beschriebenen Forschungsdesigns, welches einer Mixed-Methods-Herangehensweise entspricht, wird in einer quantitativ orientierten Studie versucht, die Ergebnisse der qualitativen Studien aus Kapitel 5 und der bestehenden Literatur zu bestätigen, respektive zu falsifizieren. Der objektive Charakter der Clickstream-Daten erlaubt eine exakte Beschreibung des Nutzerverhaltens ohne subjektive Einflüsse oder Effekte, die bei qualitative Methoden wie Interviews oder Fragebögen-Umfragen auftreten können. Die folgenden fünf Leitfragen sind beeinflusst durch die Ergebnisse der Twitter-Umfrage und sollen die in Abschnitt 5.1 formulierten Leitfragen der Twitter-Umfrage konkretisieren.

### Wodurch ist Twitter-Nutzung charakterisiert?

Unabhängig von der primären Fragestellung und den Leitfragen, die mit Wiederfinden in Zusammenhang stehen, existieren auch Fragen, die die generelle Twitter-Nutzung betreffen. Es ist wenig darüber bekannt, wie Nutzer ihre Tweets lesen bzw. ihre *Timeline* konsumieren. Lediglich die Arbeit von Counts und Fisher, beschrieben in Abschnitt 3.2.1.1, gibt hier erste Einblicke (Counts & Fisher, 2011). Mögliche offene Fragen dabei sind: Wie häufig gehen Nutzer online? Wie viele Twitter-Sitzungen hat ein Nutzer pro Tag? Wie lange dauern Twitter-Sitzungen? Zu welcher Tageszeit findet die meiste Interaktion statt? Welche Unterseiten von Twitter werden häufig besucht? Diese und weitere Fragen werden zu Beginn in Abschnitt 6.4.3 der Arbeit behandelt und zeigen, in welchem Verhaltenskontext das Wiederfindensverhalten eingebettet ist.

### Wie kann Wiederfinden in Clickstream-Daten identifiziert werden?

In der Literatur werden zahlreiche Verfahren verwendet, um Re-finding-Verhalten in Logs zu identifizieren. Diese beziehen sich auf andere Kontexte wie Wiederfinden von E-Mails oder im Web. Sie werden in Abschnitt 6.4.4 kurz erläutert. Wie Re-finding in Social-Media-Clickstream-Daten identifiziert werden kann bzw. welche Auswirkungen unterschiedliche Ansätze auf Erkennungsgütekriterien wie *recall* und *precision* besitzen, ist nicht erforscht. Dabei haben diese Ansätze starke Auswirkungen auf Wiederfindenscharakteristika, wie die Anzahl der wiederfindenden Personen. Die Diskussion um die Identifikation von Wiederfindenssitzungen ist zentral für die Fragestellungen in dieser Arbeit.

### Wie häufig werden Tweets wiedergefunden?

Da bis dato keine quantitativen Informationen darüber vorliegen, ob und in welchem Ausmaß Tweets wiedergefunden werden, steht zunächst die Frage im Vordergrund, wie häufig Wiederfinden tatsächlich stattfindet und wie viele Nutzer in Wiederfindensaktionen involviert sind. Wie häufig tritt Wiederfinden generell auf? Gibt es Tweets, die häufiger wiedergefunden werden als nur einmal? Welcher Tweet wird am häufigsten wiedergefunden? Wie hoch ist die *message uncertainty ratio*<sup>2</sup> bei Wiederfindensaktionen (Elsweiler, Harvey & Hacker, 2011)? All dies sind grundsätzliche Fragen, die für andere Medien und Domänen wie Websuche oder E-Mailsuche bereits in mehrere Arbeiten untersucht wurden (vgl. Abschnitt 2.2.4) und für Wiederfinden auf Twitter im Rahmen der Logdatenanalyse geklärt werden müssen. Zusätzlich stellt diese Frage eine Konkretisierung einer ähnlichen Frage aus den Leitfragen zur Twitter-Umfrage dar. Die dort gegebenen, subjektiven Antworten in Bezug auf Häufigkeit von Aufbewahren und Wiederfinden von Tweets sollen mit den Daten in der lebensnahen Nutzungssituation der Logstudie kontrastiert werden.

### Wodurch ist der Lebenszyklus eines Tweets charakterisiert?

<sup>2</sup>*message uncertainty ratio*  $MUR = \text{GeklickteTweets} / \text{EinzelnegeklickteTweets}$ . Erfolgen in einer Sitzung sechs Klicks auf vier unterschiedliche Tweets, beträgt die  $MUR = \frac{6}{4} = 1,5$

Der *Lebenszyklus* eines Tweets thematisiert die Frage nach der zeitlichen Dimension des Wiederfindens, ergo der Zeitperiode zwischen erstem Kontakt und erneutem Wiederfinden. Wie lange kann diese Zeitspanne sein? Welche ist die durchschnittliche Zeit, in der Tweets wiedergefunden werden? Ferner stellt sich die Frage, ob das Alter eines Tweets beim ersten Kontakt wichtig für dessen Lebenszyklus ist. Beides steht im Zusammenhang mit der Frage: Wie lange sind Tweets von Relevanz für einen Nutzer? Untersuchungen, die sich mit der Relevanz von Tweets auf einer allgemeinen, globalen Ebene auseinandersetzen und analysieren, welche Faktoren zur Verbreitung von Tweets auf der Plattform beitragen, wurden bereits in Abschnitt 3.2.1.3 dargestellt und dort im Kontext des globalen Lebenszyklus vorgestellt. Für das PIM-Verhalten von Nutzern ist der nutzerbezogene, individuelle Lebenszyklus eines Tweets von Bedeutung. Weitere Bezüge zwischen PIM und Lebensdauer eines Tweets, könnten hier auch zur Dauer bis Tweets gelöscht werden, hergestellt werden. Wichtig dabei ist auch die Frage, wie Twitter-Funktionen (Favorisieren/*liken* und RT/Weiterleiten) den individuellen Lebenszyklus von Tweets beeinflussen.

#### **Welche Wiederfindensstrategien verwenden Nutzer?**

Eine der wichtigsten Leitfragen ist sicherlich die Frage nach dem Wiederfindensverhalten. Welche Strategien wenden Nutzer an, um zu vorher gesehenen Tweets zurückzukehren? Wie viel Zeit verbringen Nutzer mit Wiederfinden? Gibt es Interaktionsketten, die Wiederfinden charakterisieren? Welche Twitter-internen Funktionen verwenden Probanden hierfür? Können Verhaltensweisen wie gezielte Schlüsselwortsuche (*teleporting*) und Navigation (*orienting*), die aus Studien zum Wiederfindensverhalten in anderen Domänen bekannt sind, auch beim Wiederfinden von Social-Media-Nachrichten identifiziert werden (Barreau & Nardi, 1995; Teevan et al., 2004)? Haben unterschiedliche Personen unterschiedliche Wiederfindensstrategien? In engem Zusammenhang mit Wiederfindensstrategien steht natürlich auch das Aufbewahrensverhalten der Nutzer, da manche Wiederfindensstrategien vorheriges Aufbewahren voraussetzen. Bewahren Nutzer Tweets auf, um diese später wiederzufinden? Welche Twitter-Funktionen werden zum Aufbewahren von Tweets genutzt? Stehen Wiederfindensstrategien mit Aufschiebestrategien im Einklang? Wie gut sind Nutzer bei der Vorhersage zukünftiger Informationsbedürfnisse (Bruce, 2005)? Können Faktoren ausgemacht werden, die Wiederfinden als schwierige Aufgabe charakterisieren? Besteht ein Zusammenhang zwischen Wiederfindensstrategien und Tweet-Charakteristika? Abschnitt 5.6 hat gezeigt, dass das Setzen eines Lesezeichens zum Zwecke des separate Abspeicherns eines Tweets in der Favoritenliste einer der Hauptmotivationen ist, Tweets zu favorisieren. Jedoch konnte auch festgestellt werden, dass die Favoritenliste nur selten besucht wird, was letztendlich die Frage nach den Gründen für diese Diskrepanz aufwirft. Mögliche Erklärungsansätze liefern die in Abschnitt 2.3 vorgestellten Hypothesen zum PIM-Verhalten.

#### **Welche Charakteristika besitzen wiedergefundene Tweets?**

In engem Zusammenhang mit der Frage aus welchen Gründen Nutzer Tweets wiederfinden wollen und was sie dazu motiviert, steht auch die Frage nach den Charakteristika der wiedergefundenen Tweets. Besitzen wiedergefundene Tweets immer einen Link, ein Bild oder Video? Sind es bestimmte Hashtags oder *@Mentions*, die Nutzer zum Wiederfinden motivieren? Werden längere Tweets eher wiedergefunden als kürzere Tweets? All diese Fragen sind von besonderer Relevanz und werden in Abschnitt 6.4.5.2 indirekt in Form eines Vorhersageexperiments, in welchem diese Eigenschaften als Features verwendet werden, behandelt.

## **6.2. Logdatenanalyse als Methode**

Der folgende Abschnitt zeigt zunächst die Bedeutung der Logdatenanalyse für das Fach der Informationswissenschaft bzw. der *Library and Information Science* (LIS) auf. Anschließend wird eine

Charakterisierung von Logdaten und ihren unterschiedlichen Varianten und Bezeichnungen vorgenommen. Ferner wird die Methodik der Logdatenanalyse, ihre Vorteile sowie ihre Einschränkungen diskutiert. Abschließend werden die Phasen der Logdatenanalyse sowie relevante Arbeiten für diese Phasen vorgestellt.

### 6.2.1. Logdatenanalyse im informationswissenschaftlichen Forschungskontext

Widmet man sich der Frage, welche Rolle die Logdatenanalyse bei der Untersuchung des Informationsverhaltens einnimmt, lohnt es sich einerseits Meta-Studien heranzuziehen, welche sich damit beschäftigen, welche Forschungsstrategien sowie Datensammlungs- und Analysemethoden eingesetzt werden und andererseits Sekundärliteratur zu sichten, welche einen inhaltlichen und methodischen Überblick über das Fach gibt. Somit ergänzt dieser Abschnitt die konzeptuelle Diskussion der Fragestellung was Informationsverhaltensforschung ausmacht, — durchgeführt in der Einleitung des Kapitels 2.1 — um eine empirisch, methodische Perspektive.

Grundsätzlich stellt sich die Frage, ob die Methode der Logdatenanalyse eine Methode ist, die zur Untersuchung des Informationssuchverhaltens, respektive für Studien, die auf die Erforschung von „information needs, uses, seeking and sense-making“ (Case, 2007, S.192) abzielen, erprobt ist. Ein Leitwerk der Sekundärliteratur zur Informationssuchverhaltensforschung — und aus der Methodennutzung betrachtet das einzige relevante — stellt das Buch „Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior“ von Case und Givens dar (Case & Given, 2016). Case und Given erwähnen die Logdatenanalyse nur in einem Nebensatz und geben ihr bei der Vorstellung von Methoden zur Untersuchung des Informationsverhaltens keinen weiteren Raum (Case & Given, 2016, S.254). Interessanterweise zitieren sie die Arbeit von Rice und Borgman, welche für den Einsatz von sogenannten *computer-monitored data* bei informationswissenschaftlicher Forschung plädieren und IR-Systeme als Beispielsysteme anführen, bei denen derartige Studien durchgeführt werden können (Rice & Borgman, 1983, S.247).

Meta-Studien zur Methodenanalyse finden sich sowohl für den engeren Kontext der ISV-Forschung als auch für den breiteren Kontext der Bibliotheks- und Informationswissenschaft (*Library and Information Science LIS*), in welche *Information Seeking* als derjenige Teilbereich eingeordnet werden kann, der das Verhalten von Nutzern beim Umgang mit Informationssystemen abdeckt. Als Meta-Studien in der ISV-Forschung sind vor allem die Arbeiten von Heidi Julien et al. aus den Jahren 1996, 2000, 2011 und 2014 zu nennen (Julien, 1996; Julien & Duggan, 2000; Julien et al., 2011; Julien & O'Brien, 2014). Die Studie aus dem Jahr 2014 ist die aktuellste Ausgabe einer Langzeitanalyse zur ISV-Forschung und fasst die Ergebnisse der älteren Studien zusammen, so dass ein Trend in der Methodennutzung über 30 Jahre hinweg dargestellt wird (Julien & O'Brien, 2014).

Julien und Kollegen können zeigen, dass der Anteil an Studien, die die Methode der Logdatenanalyse zur Untersuchung des Informationssuchverhaltens verwenden, über die Jahre hinweg von 4% (1996) auf 12,4% (2000) steigt und schließlich auf 9,8% im Jahr 2011 bzw. 6% im Jahr 2014 fällt (Julien & O'Brien, 2014). Es ist anzumerken, dass einige Studien auch als *Mixed Methods* klassifiziert werden, wobei nicht erkenntlich ist, ob ein gewisser Teil dieser Studien auch Logdaten auswertet. Tuomaala, Järvelin und Vakkari merken an, dass der Anteil an Logstudien bei Meta-Studien jedoch kritisch zu reflektieren ist, da häufig nur Journal-Artikel untersucht werden, viel informationswissenschaftliche Forschung aber auch auf Konferenzen in Form von *Proceedings* veröffentlicht wird (Tuomaala et al., 2014, S.1450). Eine vergleichbare Analyse legt Greifender vor (Greifeneder, 2014). Sie untersucht 155 ISV-Studien im Zeitraum von 2012 bis 2014 aus den Quellen *JASIST*, *Information Research*, *The Journal of Documentation* und den Veröffentlichungen der

iConference. Die Artikel werden manuell durch Browsing auf Schlüsselwörter hin durchsucht bzw. ausgewählt, wodurch eine gewisse Subjektivität nicht auszuschließen ist. In den ausgewählten Artikeln kommt die Logdatenanalyse nicht häufig genug vor, weshalb sie unter die Kategorie *Other* fällt.

Im breiteren Kontext der LIS finden sich ebenfalls Meta-Studien zur Methodennutzung. Koufogiannakis et al. untersuchten 2664 Journal Artikel von 91 LIS-Journals aus dem Jahr 2001 (Koufogiannakis et al., 2004). In ihrer Untersuchung wird jedoch die Variable *Untersuchungsmethodik* nicht erhoben bzw. diskutiert. Hider und Pymm weisen in ihrer Inhaltsanalyse von 561 Artikeln aus 20 LIS-Journals darauf hin, dass neuere Strategien und Techniken wie die Logdatenanalyse in vorherigen Untersuchungen zur Methodennutzung unterrepräsentiert sind bzw. nicht in entsprechenden Taxonomien geführt sind und geben die Empfehlung, die Logdatenanalyse in zukünftigen Studien als gesonderte Methode zu führen. In ihrer Untersuchung wird sie jedoch noch unter *other strategies* geführt (Hider & Pymm, 2008). Eine Studie, die drei Jahrzehnte (1965-1995) an Methodennutzung in 37 LIS Journals überblickt, ist die Arbeit von Järvelin und Vakkari aus dem Jahr 1993 (Järvelin & Vakkari, 1993). Ähnlich wie bei den Arbeiten von Julien und Kollegen wird diese Arbeit von Tuomaala, Järvelin und Vakkari erneut aufgegriffen, um eine Langzeitanalyse durchzuführen (Tuomaala et al., 2014). Erstaunlich dabei ist, dass sowohl in den früheren Arbeiten als auch in der Ergänzung im Jahr 2014 die Methodik der Logdatenanalyse nicht aufgeführt wird. Dies liegt wohl weniger an der Tatsache, dass diese nicht zum Einsatz kommt, als daran, dass für die entsprechenden Variablen (*research strategies, data-collection methods, types of analysis*) die Ausprägung Logdatenanalyse nicht vorhanden war, und diese vermutlich unter anderen Ausprägungen subsumiert wurde (Tuomaala et al., 2014, S.1452). Eine erst kürzlich erschienene Arbeit von Risso, die ebenfalls die wissenschaftliche Methodennutzung in der LIS diskutiert, führt die Logdatenanalyse nicht auf (Risso, 2016). Methodisch verwenden die bis dato erwähnten Arbeiten alle qualitative Methoden, wie die Inhaltsanalyse. Aber auch Arbeiten, die automatisierte Verfahren zur Themenentdeckung auf Korpora von LIS-Arbeiten anwenden, können keine Spuren von Logdatenanalyse als Methodik vorweisen (Sugimoto et al., 2011).

Parallel zu den oben genannten Arbeiten existiert eine Reihe bereits älterer Publikationen, die schon in frühen Jahren einen Überblick über die Nutzung der Logdatenanalyse bei der Untersuchung des Verhaltens von Nutzern im Bibliothekskontext beim Umgang mit Online-Katalogsystemen geben. Thomas A. Peters veröffentlicht 1993 einen Artikel mit dem Titel *The History and Development of Transaction Log Analysis*, in welchem die Geschichte und Entwicklung der Logdatenanalyse der ersten 25 Jahre vorgestellt wird (Peters, 1993). In diesem Artikel weist Peters darauf hin, dass bereits in früheren Jahren Literaturauswertungen zur Methode der *Transaction Log Analysis* durchgeführt wurden und zählt in chronologischer Reihenfolge die Arbeiten von Hansen (1980), Borgman und Rice (1982), Cochrane und Markey (1983), Borgman (1986), Lewis (1987), sowie Simpson und Larson (1991) auf (Hansen, 1980; Borgman & Rice, 1982; Cochrane & Markey, 1983; Borgman, 1986; Lewis, 1987; Simpson & Larson, 1991). In den Folgejahren — parallel zur Ankunft der ersten Websuchmaschinen — erscheinen weitere Überblicksarbeiten, die die Erkenntnisse verschiedener Logdatenstudien zusammenfassen. Jansen und Spink nehmen eine vergleichende Analyse von neun Studien zur Nutzung von Websuchmaschinen vor, indem sie Trends und Unterschiede bei verschiedenen Charakteristika der Suche — wie Ein-Term-Suchanfragen — kontrastieren (Jansen & Spink, 2006). In ihrem Artikel *Twenty-Five Years of End-User Searching, Part 1: Research Findings* nimmt Markey eine ausführliche Meta-Analyse zahlreicher Logdatenstudien vor und untersucht diese hinsichtlich bestimmter Aspekte wie *durchschnittliche Querylänge* oder *Nutzung von Suchoperatoren* (Markey, 2007). Agosti, Crivellari und Di Nunzio geben einen Überblick über die letzten zehn Jahre an Logdatenuntersuchungen mit Websuchmaschinen und digitalen Bibliothekssystemen (Agosti et al., 2011). Hierfür analysieren sie 50 Paper und Poster aus den

Datenbank	Einträge	Suchanfrage	Ergebnis
The ACM Full-Text Collection	446.496	(clickthrough clickstream „search logs“ „query logs“)	921
The ACM Guide to Computing Literature	2.578.290	(clickthrough clickstream „search logs“ „query logs“)	1.593
IEEE Xplore Digital Library	4.016.518	(((((clickstream) OR clickthrough) OR „query logs“) OR „search logs“) TOPIC:(clickthrough)	1.903
Web of Science All Databases	~ 1 Mrd.	OR TOPIC:(clickstream) OR TOPIC: („query logs“) OR TOPIC: („search logs“) clickstream in All Fields OR clickthrough in All Fields	452 (123)
Wiley Online Library	~ 6 Mio.	OR „query logs“ in All Fields OR „search logs“ in All Fields	614

**Tabelle 6.1.:** Suche in bibliographischen Datenbanken nach Literatur zum Thema Logdatenanalyse

Veröffentlichungen von fünf großen Information-Retrieval-Konferenzen<sup>3</sup> und 30 Journalartikeln (Agosti et al., 2011).

Schließlich kann die Bedeutung der Logdatenanalyse auch anhand einer stichprobenartigen Suche in verschiedenen, für die Informationswissenschaft bedeutsamen (bibliographischen), Literaturdatenbanken abgelesen werden. Tabelle 6.1 zeigt die Datenbanken, die für diesen Zweck verwendet wurden, welche Suchanfrage formuliert wurden, sowie die Höhe der Suchergebnisse für die jeweilige Datenbank. Für die Suchanfrage wurden zusätzlich Terme verwendet, die als Synonym zu *Logs* interpretiert werden können (vgl. Abschnitt 6.2.2). Als bester Referenzpunkt kann die Suche in der *ACM Full-Text Collection* betrachtet werden, da diese die Tagungsbände der wichtigsten Konferenzen im IR und ISV beinhaltet. Eine große Überschneidung besteht bei den Ergebnissen der Suche in den Datenbanken *ACM Guide to Computing Literature* und der *IEEE Xplore Digital Library* da erstere auch bibliographische Daten von Veröffentlichungen anderer Verlage beinhaltet. Bei einer stichprobenartigen Untersuchung der jeweiligen Treffermengen konnte festgestellt werden, dass die Trefferliste der *IEEE Xplore Digital Library* einige falsch Positive enthielt. Die Suche in der Datenbank *Web of Science* besitzt zusätzlich eine Facettenklassifikation aus der abgelesen werden konnte, dass 123 der 452 Ergebnisse der Kategorie *Research Area: Information Science/Library Science* zugeordnet werden und in diesem Kontext einschlägig sind. Insgesamt betrachtet zeigen alle Ergebnisse eine solide Anzahl an Veröffentlichungen (Konferenzbeiträge, Journalartikel, Bücher), die Logdatenerhebung und -analyse thematisieren.

Eine ganzheitliche Auflistung und Diskussion sämtlicher Meta-Studien im Informationsverhalten oder LIS-Kontext und eine detaillierte Analyse aller Logdatenstudien im Information Retrieval-Kontext liegt außerhalb des Rahmens dieser Arbeit. Jedoch ergeben sich für die Nutzung der Methodik in dieser Arbeit folgende Erkenntnisse: Zahlreiche Arbeiten wie die von Peters, Markey oder Jansen und Spink sowie die Veröffentlichungszahlen zum Themenbereich *Logdaten* in bibliographischen Datenbanken zeigen, dass die Logdatenanalyse eine etablierte Methodik ist, die häufig diskutiert und genutzt wird (Peters, 1993; Markey, 2007; Jansen & Spink, 2006). Es ist überraschend, dass eine Methodik, die offensichtlich durchaus etabliert ist, in verschiedenen LIS-Überblicksarbeiten nicht berücksichtigt oder erwähnt wird. Es scheint fast so, dass Arbeiten, die nicht aus einer gewissen wissenschaftlichen Disziplin stammen und eine entsprechend qualitative Herangehensweise an die Analyse des Informationssuchverhaltens besitzen, deswegen auch nicht berücksichtigt werden. Eine Tatsache, die vor allem auch bei Donalds Case Buch zu erkennen

<sup>3</sup>SIGIR, ECIR, WWWW, JCDL, ECDL



ist. Erst kürzlich ist hier die vierte Auflage erschienen (jetzt mit Lisa Given als Co-Autorin) und, obwohl Websuchmaschinen nun schon seit Jahrzehnten etablierte Informationssysteme mit Milliarden von Suchanfragen pro Tag<sup>4</sup> sind, finden sich in Case und Givens' Buch noch immer keine Forschungsarbeiten von Eric Horvitz, Jaimie Teevan, Ryan W. White, Susan Dumais und anderen, welche Logdatenanalyse als Methodik benutzen, um detailliert die Informationsbedürfnisse und das Suchverhalten von Menschen bei der Websuche zu analysieren (Case & Given, 2016). Case und Given argumentieren hier wie folgt: „Studies of people's system searching practices (i.e. what are classically referred to as 'information retrieval' studies) are not defined as information behaviour studies“ (Case & Given, 2016, S.15). Für Case und Given sind diese Art von Studien zu wenig auf die Person- bzw. den Kontext fokussiert. In der vorliegenden Arbeit werden derartige Studien durchaus als Informationsverhaltensstudien charakterisiert, da sie eine Brücke zwischen Nutzerverhalten und Systemverbesserung darstellen. Ferner sollten diese Arbeiten als Informationsverhaltensarbeiten charakterisiert werden, da die Methodik der Logdatenanalyse *a priori* als Methode gedacht ist, um das Informationsverhalten oder zumindest das Nutzungsverhalten von Personen zu erheben und näher zu beschreiben, wie Abschnitt 6.2.3 zeigen wird. Es scheint also ein gewisser Dissens zwischen einer Methodik zu bestehen, die seit jeher als Methode zur Untersuchung des Informationssuchverhaltens gedacht war und einer wissenschaftlichen Gemeinschaft, die selten auf diese Methode zurückgreift. Im Folgenden wird die Logdatenanalyse methodologisch im Detail charakterisiert.

### 6.2.2. Charakterisierung von Logdaten

Spuren des menschlichen Verhaltens — im Englischen auch als *behavioral logs* bezeichnet — werden in der sozialwissenschaftlichen Forschung schon lange als Datengrundlage für Studien genutzt (Bortz & Doering, 2006). Dort werden beispielsweise Eselsohren in Büchern oder stark abgenutzte Teppichböden in Museen als Spuren menschlichen Verhaltens für Untersuchungen verwendet. Auch beim Umgang mit Computern hinterlassen Nutzer Spuren im virtuellen Raum, die in Daten resultieren und zur Untersuchung von Verhalten gespeichert (*logging*) und ausgewertet werden können. Jansen, Taksa und Spink verwenden hierfür den Begriff der *transaction logs* den sie wie folgt definieren: „[Transaction Log Analysis is] an electronic record of interactions that have occurred between a system and the user of a system“ (Jansen et al., 2009, S.2). Für sie ist *transaction log analysis* (TLA) eine breite Kategorie an Verfahren, welche Unterkategorien enthalten kann, wie *web log analysis*, also die Analyse von HTTP GET/POST Anfragen (*requests*) an einen Server oder die Analyse von Suchanfragen bei Suchmaschinen, welche als *search log analysis* oder *query log analysis* bezeichnet wird (Jansen et al., 2009, S.2). Peters definiert Logdaten ähnlich, fokussiert aber in seiner Definition explizit auf Information-Retrieval-Systeme bzw. Online-Katalog-Systeme von Bibliotheken (Peters, 1993). Dumais und Kollegen besitzen einen weitgefassten Logdaten-Begriff: „In HCI research behavioral logs arise from the activities recored when people interact with computer systems and services“ (Dumais et al., 2014, S.350). Als Beispiel führen sie folgenden Aktivitäten an:

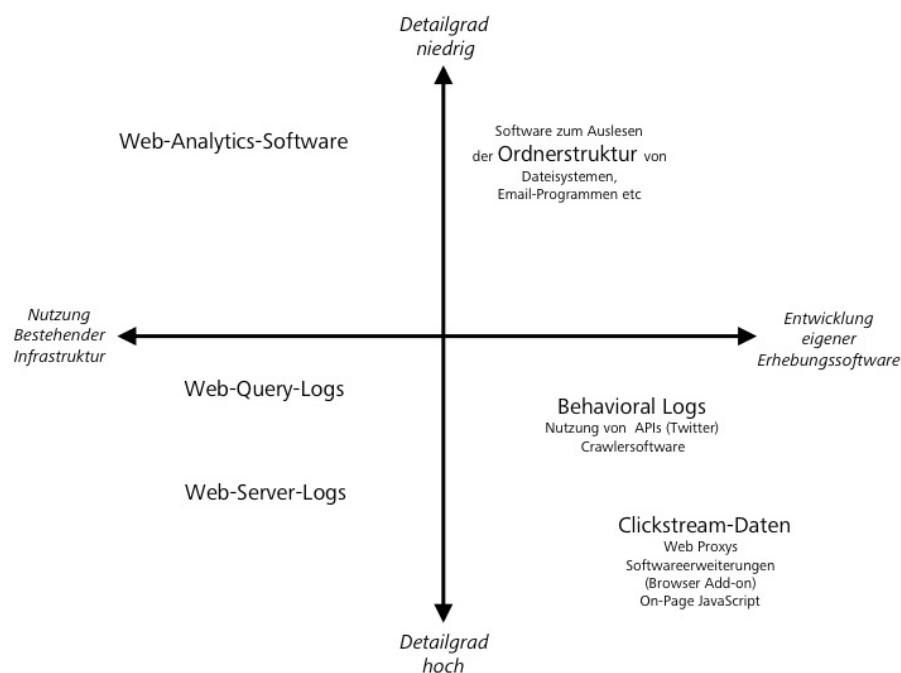
- Low-level-Aktionen wie das Drücken von Tasten
- Inhalte, die im Browser betrachtet werden
- Suchanfragen, die an Websuchmaschinen gerichtet werden
- Verhaltensmuster beim Durchstöbern von E-Commerce-Seiten

Darüber hinaus werden von Dumais et al. auch physiologische Daten wie Eye-Tracking-Daten zu Logdaten gerechnet. Explizit werden ebenfalls Veränderungen bei Wikipedia-Artikeln, Social-

<sup>4</sup>Laut Artikel der Daily Mail vom 28. Juni 2016 sind es 54 907 Suchanfrage pro Sekunde (<http://dailymail.ai/2941x49>)

Media-Nachrichten wie Tweets oder Fotos auf Flickr als frei zugängliche *behavioural logs* gewertet (Dumais et al., 2014, S.350). Lazar, Fang und Hochheiser werten auch erstellte Ordnerstrukturen in Dateisystemen, E-Mail-Programmen sowie Web-Bookmarks als Logdaten, da die Anzahl an Ordner oder die Tiefe der Ordnerstruktur Aufschluss darüber geben, wie Nutzer mit Anwendungen interagieren und Informationen speichern und verwalten (Lazar et al., 2010, S.315f.). Die Definitionen und Beispiele zeigen, dass sich Logdaten in mehreren Dimensionen unterscheiden können und dadurch auch unter unterschiedliche Begrifflichkeiten firmieren. Jegliche Interaktion zwischen Nutzern und einem Informationssystem können als *Logs* aufgezeichnet werden, wobei im Kontext des Informationssuchverhaltens oder des Information Retrieval Suchmaschinenlogdaten gemeint sind.

Abbildung 6.1 verortet mögliche Bezeichnungen bzw. Erhebungsverfahren von Logdaten in einem zweidimensionalen Koordinatensystem mit den Achsen *Detailgrad der Logdaten* sowie *Art der Erhebung*. Durch das Koordinatensystem soll verdeutlicht werden, dass ein gewisser *Trade-off* besteht zwischen serverseitiger Datensammlung, welche nicht alle Nutzer-System-Interaktionen enthalten kann und client-seitigen Datensammlungsmethoden durch adaptierte Software, die sehr detaillierte Clickstream-Daten erheben kann, aber unter Umständen sehr viel Entwicklungsaufwand in sich birgt. Lazar, Fang und Hochheiser bezeichnen diese Adaption auch als *instrumented software*. Zahlreiche Studien aus Abschnitt 2.2.4.2 arbeiten mit Varianten von veränderten Browsern, die unter diese Kategorie fallen (Cockburn et al., 2003; Obendorf et al., 2007; Sanderson & Dumais, 2007). Vergleichbare Ansätze für clientseitige Erhebung stellen Browsererweiterungen dar oder ein auf der jeweiligen Website eingebettetes JavaScript.



**Abbildung 6.1.:** Verortung unterschiedlicher Logdatenerhebungsmethoden in einem Koordinatensystem mit den Achsen *Detailgrad der Logdaten* und *Entwicklungsaufwand*

Logdatenuntersuchungen können — unabhängig vom Detailgrad der Daten — unterschiedliche Ziele verfolgen. Zum einen kann es das Ziel sein, systemseitige Verbesserungen vorzunehmen, indem Logdaten zur Überwachung der Systemperformance (Error-Logs) für Debuggingzwecke oder dem Aufdecken von Usability-Problemen einer Website genutzt werden (Lazar et al., 2010, S.310). Zum anderen werden Logdaten dazu eingesetzt, um ein besseres Verständnis darüber zu bekommen, wie Menschen Systeme nutzen. Dies ist ein Ziel, das bereits in Abschnitt 6.2.1 thematisiert wurde. Zum Beispiel können Suchanfragelogs (*query logs*) ausgewertet werden, um das *Information Seeking Behaviour* des Nutzers zu verstehen. Auch hier kann dieses Verständnis zu einem späteren Zeitpunkt in Systemverbesserungen münden (Meister & Sullivan, 1967; Penniman, 2009). Penniman formuliert dies wie folgt: „The promise of analyzing transaction logs has always been at least twofold: first to *describe* what users actually do while interacting with a system and second to use this understanding to predict what should be the next actions they might take to use the system effectively“ (Penniman, 2009, S.19). Penniman beschreibt die Vision eines intelligenten Assistenzsystems, das — basierend auf Erkenntnissen einer Logdatenuntersuchung — dazu in der Lage ist, den Nutzer durch die Vorhersage seiner zukünftigen Schritte zu unterstützen. Support bei der Suche ist ebenfalls ein Ziel, das bei Jansen aufgeführt wird (Jansen, 2006). Dumais und Kollegen unterscheiden zwei Arten von Logdatenstudien: *observational log studies*, deren Intention es ist, menschliches Verhalten durch Logdatenanalyse zu verstehen und *experimental log studies*, die das Ziel verfolgen, alternative Systemeinstellungen z. B. in Form eines A/B-Tests zu vergleichen (Dumais et al., 2014).

### 6.2.3. Methodische Charakterisierung der Logdatenanalyse

Die ersten Überlegungen zur Nutzung der *transaction log analysis* als wissenschaftliche Methode zur Beobachtung des menschlichen Verhaltens beginnt Mitte der 1960er-Jahre durch die Psychologen Edwin Parker und William Paisley (Peters, 1993, S.42). Sie argumentierten, dass schon in den frühen Jahren der TLA die Evaluation eines IR-Systems stets auch verhaltensorientierte Kriterien seitens des Nutzers betrifft, und einer Evaluation nicht ausschließlich technologische Kriterien zu Grunde liegen (Parker & Paisley, 1966). Interessanterweise wird die Auswertung von Logdaten seit ihrer frühen Phase nicht nur als Möglichkeit gesehen, um Systeme zu evaluieren, sondern sie war *a priori* auch dazu gedacht, um mehr über das Verhalten von Nutzern bei der Interaktion mit dem System zu lernen (Parker & Paisley, 1966; Robertson & Hancock-Beaulieu, 1992). Seit ihres Bestehens ist sie folglich eine Methodik der Informationsverhaltensforschung. Die methodologische Diskussion um die TLA setzt sich während der 1980er Jahre fort und wird unter anderem von Forschern wie Cochrane und Markey, Kase und Sanders sowie Rice und Borgman geführt (Peters, 1993). Robertson und Hancock-Beaulieu beschreiben wie die Entwicklung der Information-Retrieval-Forschung hin zum sogenannten *Interactive Information Retrieval* die Schwierigkeit und Komplexität von Experimenten erhöht, sehen aber gleichzeitig die Chance, die durch Methoden wie der TLA einhergehen, um das Wissen über das menschliche Informationssuchverhalten zu erhöhen (Robertson & Hancock-Beaulieu, 1992). Eine theoretische und methodologische Auseinandersetzung mit der Logdatenanalyse nehmen Jansen, Taksa und Spink vor (Jansen et al., 2009). Nach Jansen, Taksa und Spink sowie auch Dumais und Kollegen fällt die Logdatenanalyse in das behavioristische Paradigma von Wissenschaft, da der Fokus auf der deskriptiven Beobachtung und dem Festhalten von sich tatsächlich ereignetem Verhalten liegt, wobei im vorliegenden Kontext virtuelle Spuren als Indikatoren für menschliches Verhalten gewertet werden (Jansen et al., 2009, S.4) (Dumais et al., 2014, S.349). Methodisch gehört die Logdatenanalyse zur Kategorie der nonreaktiven Erhebungsmethoden, da sie eine Möglichkeit zur Sammlung von Interaktionsinformationen zwischen System und Nutzer bietet, ohne eine explizite Einbindung des Nutzers — beispielsweise in Form einer experimentellen Situation — vornehmen zu müssen (Jansen et al., 2009, S.2) (Bortz

& Doering, 2006, S.710). Logdatenerhebungen kennzeichnen sich vor allem dadurch, dass eine lebensnahe Datenerhebung möglich ist, wobei eine Aufzeichnung von Benutzeraktivitäten *in situ* durchgeführt werden kann, ohne dass ein Eingriff in das natürliche Umfeld der Versuchsperson erfolgen muss (Dumais et al., 2014, S.350). Es wird keine experimentelle Situation hergestellt, weshalb auch kein Einfluss auf die Versuchsperson, die Ereignisse oder die Prozesse erfolgt (Bortz & Doering, 2006, S.326). Es wird tatsächliches Verhalten beobachtet, so dass letztendlich nicht auf vergangenes Verhalten geschlossen werden muss, wie bei klassischen nonreaktiven Methoden. Praktisch umfasst die Logdatenanalyse drei Phasen: die Datensammlung, die Datenaufbereitung und die Datenanalyse (Jansen, 2006, S.407). Diese Phasen können je nach Erhebungskontext von unterschiedlich hohem Arbeitsaufwand geprägt sein, was auch in den Vorteilen und Nachteilen der Methodik erkenntlich wird, welche im Folgenden aufgeführt werden:

Jansen, Taksa und Spink verweisen auf die Vorteile von Logstudien im Kontrast zu Laborexperimenten in den Dimensionen: Umfang, Mächtigkeit, Fokussierung, Ort und Dauer (Jansen et al., 2009, S.10).

- In Abhängigkeit zur Erhebungsform können Logdatenstudien sehr effizient sein, da Daten genutzt werden können, die ohnehin z. B. als HTTP-Requests am Server anfallen.
- Im Gegensatz zu Laborexperimenten besitzen Logdatenstudien keine Einschränkungen in Bezug auf die Stichprobengröße, den Ort oder die Dauer der Datenerhebung. Mit dem Umfang steigt auch die statistische Aussagekraft und selbst kleine Unterschiede zwischen Populationen können beobachtet werden (Dumais et al., 2014, S.352).
- Als nonreaktive Methoden können Nebeneffekte wie Interviewereffekt, Versuchsleiterereffekt, bewusste Testverfälschung oder Antwortverzerrung eliminiert werden (Bortz & Doering, 2006, S.325).
- Logdatenstudien werden folglich allgemein als objektiver, realistischer und repräsentativer erachtet (Höchstötter, 2007, S.136).

Eine detaillierte Auseinandersetzung mit den Einschränkungen von Logdatenstudien nimmt Kurth vor, wobei dieser explizit zwischen *limits* — also eher konzeptionellen, methodischen Grenzen — und *limitations* — also eher pragmatischen Schwierigkeiten — unterscheidet (Kurth, 1993).

- Sämtliche Formen nonreaktiver Erhebungsmethoden sind durch eine eingeschränkte Kontrolle des Datenerhebungsprozesses charakterisiert. So sind in der Regel keine soziodemographischen Informationen über die Nutzer vorhanden (Höchstötter, 2007, S.136). Generell ist die Kontrolle im Datenerhebungsprozess abhängig davon, ob bestehende Infrastruktur genutzt wird oder eigene technische Lösungen entwickelt werden (Lazar et al., 2010; Dumais et al., 2014).
- Es gibt kein randomisiertes Sampling der Studienteilnehmer, keine Kontrolle über die durchgeführten Aufgaben sowie keine Kontextinformationen über Motivation, Erfolg oder Zufriedenheit bei der Systemnutzung (Dumais et al., 2014, S.352).
- In Bezug auf *limitations* sind Logdaten häufig mit einem größeren Ausmaß an Säuberung (z. B. von Bot-Anfragen), Filterung und Aufbereitung der Daten verbunden (Jansen et al., 2009, S.10).
- Logdatenstudien sind stets mit ethischen und juristischen Fragestellungen verbunden. Grundsätzlich stellt sich die Frage, inwiefern durch die Erhebung und Nutzung für wissenschaftliche Analysen die Privatsphäre der Nutzer verletzt wird. Die Identifikation der Rentnerin Thelma Arnold in den AOL Suchanfrage-Logs konnte zeigen, dass oft eine einfache Anonymisierung der Daten nicht ausreicht (Barbaro & Zeller, 2006). Vor allem durch die immer stärker werdende Personalisierung der Suche in Websuchmaschinen bleibt Privatsphäre ein

kontinuierliches Problem im IR (Shen et al., 2007).

Abschließend ist darauf hinzuweisen, dass Logdaten nicht nur im lebensnahen Kontext erhoben werden können, sondern auch während Laborexperimenten. Arguello nutzt beispielsweise die Clickstream-Daten, welche bei einem Crowdsourcing-Experiment erhoben werden, zur Vorhersage der Aufgabenkomplexität bei der Websuche (Arguello, 2014).

#### 6.2.4. Anwendungsdimensionen und Phasen der Logdatenanalyse

Für das Informationssuchverhalten und das Information Retrieval hat sich die Logdatenanalyse über die letzten Jahre als sehr fruchtbare, von Erfolg gekrönte Methodik etabliert. Bereits in Abschnitt 2.2.4 wurden zahlreiche Studienergebnisse vorgestellt, die methodisch betrachtet auf Logdatenstudien basieren. Historisch gesehen werden erste Logdatenstudien bereits Ende der sechziger, Anfang der siebziger Jahre durchgeführt (Meister & Sullivan, 1967; Water et al., 1976; Nour, 1985). Als erste Logdatenstudie überhaupt gilt eine Studie von Meister und Sullivan, welche eine Auftragsarbeit der NASA darstellt und untersuchte, wie Wissenschaftler, Ingenieure und Bibliothekare der NASA ein IR-System (RECON) für bibliographische Einträge nutzen. In dieser Phase der TLA steht speziell die Reaktion auf die Ausfallzeiten und Antwortzeiten des Systems seitens der Nutzer im Vordergrund (Meister & Sullivan, 1967). Peters teilt in einem 1993 veröffentlichten Artikel das frühe Stadium der TLA (1965-1995) in drei Phasen ein (Peters, 1993, S.43):

1. Phase (Mitte 1960er bis Späte 1970er): Frühe Logstudien setzten den Fokus eher auf Systemperformanz, wie Dauer der Systemantwortzeit und die Reaktion der Nutzer.
2. Phase (Späte 1970er bis Mitte 1980er): In dieser Phase rückt das Interaktionsverhalten der Nutzer in den Fokus. Fragen nach der Systemnutzung wie Anzahl, Art und Reihenfolge von verwendeten Suchoptionen werden gestellt und beantwortet. Als Untersuchungsgegenstand dienen insbesondere Online-Katalogsysteme.
3. Phase (Seit Mitte 1980er): Die dritte Phase der TLA ist charakterisiert durch eine Diversifikation der Themen. Fortan bilden sich verschiedene Schwerpunkte heraus, wie Nutzergruppen oder Suchoptionen.

Aus aktueller Perspektive lassen sich diese Abschnitte noch um weitere Phasen ergänzen. Eine 4. Phase beginnt spätestens Ende der 1990er Jahre. In ihr konzentriert sich der Fokus auf Logdatenanalysen der ersten Websuchmaschinen (Altavista, Excite, Fireball, Lycos etc.) und dort im Speziellen auf Suchanfrageanalysen (*query log analysis*). Arbeiten in dieser Phase besitzen in der Regel einen spezifischen Auswertungsfokus (Höchstötter, 2007, S.138). Mögliche Auswertungsschwerpunkte sind: deskriptive Auswertungen, zeitliche Auswertungen und inhaltliche Auswertungen. Deskriptive Auswertungen fokussieren auf die Term- (z. B. Anzahl der Terme, Zeichenlänge), Query- (Verwendung von Operatoren), oder Sessionebene (Höchstötter, 2007, S.138). Wesentliche Arbeiten, die hier anzuführen sind, sind die Arbeiten von Silverstein et al. (Silverstein et al., 1999) sowie zahlreiche Arbeiten von Jansen und Spink (Jansen et al., 1998; Spink, Ozmutlu et al., 2002; Jansen & Spink, 2006). Einen Vergleich von Studien mit deskriptiven Auswertungsfokus nimmt Markey vor, indem sie Querylänge und andere Parameter von zahlreichen deskriptiven TLA-Studien gegenüberstellt (Markey, 2007). Zeitliche Auswertungen analysieren die durchschnittliche prozentuale Verteilung von Suchanfragen pro Wochentag oder im Verlauf eines Tages. Vergleichbare Analysen sind in die Kategorie der *time series analysis* einzuordnen (Y. Zhang et al., 2009). Die Frage nach Faktoren, die eine Session — also den Beginn und das Ende einer in sich geschlossenen Suchaktion definieren — charakterisieren, ist eine Frage, die von wiederkehrender Relevanz ist und methodisch gesehen mit unterschiedlichen Daten und Ansätze zu beantworten versucht wurde (Jansen et al., 2007; Ye & Wilson, 2014). In frühen Phasen wurde die Dauer von

Inaktivitätsphasen zwischen Aktionen (z. B. 30 Min. keine Aktion durch den Nutzer) oder der Suchanfragekontext als Sessiongrenzen gewertet (Silverstein et al., 1999). Inhaltliche Auswertungen nehmen eine Klassifizierung der Terme in unterschiedliche Themengebiete wie Erotik, Reise, Gesundheit, Ernährung oder Religion vor und untersuchen, wie Nutzer Websuchmaschinen zur Suche nach diesen Themen benutzen (Spink, Jansen et al., 2002; Cartright et al., 2011; West et al., 2013; Wan-Chik et al., 2013). Allmählich ist hierbei der Beginn einer 5. Phase zu vermerken: Die stärkere inhaltliche Auseinandersetzung führt weg von einer eher deskriptiv orientierten TLA hin zu einer TLA, die immer mehr auch das tatsächliche Informationssuchverhalten der Nutzer in den Vordergrund rückt. Immer stärker werden dabei auch Informationen über geklickte Ergebnisse mit in die Interpretation mit einbezogen und durch Platzhaltermaße — sogenannten *Proxies* — wird versucht, auf tatsächliches Nutzerverhalten zu schließen. Es werden nicht nur einfach deskriptive Analysen durchgeführt, sondern es wird auf deskriptiven Erkenntnissen aufgebaut, um beispielsweise komplexe statistische Modelle zu berechnen, die Nutzerverhalten vorhersagen oder algorithmische und/oder auch UI-bezogene Features zu entwickeln, die den Nutzer bei der Suche unterstützen und eine effektivere Suche ermöglichen. Agosti, Crivellari und Di Nunzio beschreiben eine Vielzahl an Möglichkeiten wie, auf der Analyse von Suchanfrage aufbauend, dem Nutzer Verbesserungsmöglichkeiten beim Suchanfragegenerierungsprozess mit Hilfe von *query suggestion* oder *query expansion* angeboten werden können (Agosti et al., 2011, S.665ff.). Als weitere Arbeiten können die Studien von Hassan et al. oder Odijk et al. angeführt werden, in denen mit Hilfe von Suchanfragelogs versucht wird, zu erkennen, ob Nutzer explorativ suchen oder Mühe haben, das Richtige zu finden und Nutzer entsprechend eine angepasste Ergebnisliste erhalten (Hassan et al., 2014; Odijk et al., 2015). Schließlich werden die Erkenntnisse dazu genutzt, um entsprechende Hilfestellungen in Form von Suchanfrageerweiterungen zu generieren. In dieser Phase wird auch die Bedeutung der Logdatenanalyse für die Untersuchung des Wiederfindensverhalten und des PIM im Allgemeinen offenkundig. Die in Abschnitt 2.2.4 näher vorgestellten Arbeiten von Teevan, Dumais und anderen zeigen dies. Wie bereits weiter oben erwähnt, ist ein weiterer Kontext, in welchem Clickstream-Daten genutzt werden, der E-Commerce-Bereich. Hier werden z. B. mit Hilfe von Cluster-Verfahren Informationen zum Browsing-Verhalten dazu genutzt, um Interessensmuster seitens der Nutzer zu erkennen, durch Empfehlungen darauf zu reagieren bzw. mögliches Kaufverhalten vorherzusagen (den Poel & Buckinx, 2005; Kim & Yum, 2011; Su & Chen, 2015; G. Liu et al., 2016).

Die 5. Phase ist ebenfalls dadurch gekennzeichnet, dass bei Analysen immer häufiger nicht nur Suchanfragen und geklickte Links zum Einsatz kommen, sondern auch das komplette Interaktionsbild (Mausbewegungen, Verweildauer auf Ergebnisseiten etc.) — in Abbildung 6.1 auch als Clickstream-Daten verortet — eingesetzt wird. Clickstream-Daten kommt im Kontext des IR eine besondere Rolle zu, da auf der Basis von Interaktionsdaten von Suchmaschinennutzern die ersten Learning-to-Rank-Verfahren zur automatischen Optimierung der Retrievalqualität von Suchmaschinen mit Hilfe von maschinellen Lernverfahren vorgestellt und diskutiert werden (Joachims, 2002; Agichtein et al., 2006). Seit einigen Jahren werden im *online learning to rank* im Rahmen der sogenannten *living labs for IR evaluation (LL4IR)* auch Experimente *in situ* mit Clickstream-Daten von Usern durchgeführt, die in Echtzeit bei der Interaktion mit Websuchmaschinen im Alltag erhoben werden (Schuth et al., 2015).

Im Folgenden werden Arbeiten vorgestellt, die für den Kontext der Twitter-Logstudie relevante Studien darstellen und Clickstream-Daten auf Social-Media-Plattformen erheben und verarbeiten.

### 6.2.5. Überblick zu Studien mit Social-Media-Clickstream-Daten

Social-Media-Plattformen bieten häufig API-Schnittstellen mit deren Hilfe *behavioural logs* in Form von Nutzerinteraktionsdaten abgerufen und gespeichert werden können (Dumais et al., 2014). Zahlreiche Forschungsprojekte und Studien in den unterschiedlichsten wissenschaftlichen Disziplinen greifen auf diese Methode der Logdatenerhebung zurück. Viele Studien dieser Art wurden in Kapitel 3 präsentiert. Im Folgenden liegt der Fokus nicht auf Studien, die mit gecrawlten Daten arbeiten — unabhängig davon, ob die Quelle eine offene API ist oder ein eigener Crawler dafür implementiert wurde — sondern auf Studien, die das vollständige, lebensnahe Interaktionsverhalten in Form von Clickstream-Daten verwenden, um das Nutzungsverhalten von Usern auf online-basierten, sozialen Netzwerken zu analysieren. Tatsächlich existieren nur wenige Studien, die eine Analyse von sozialen Netzwerken auf der Basis von Clickstream-Daten vornehmen. Jin et al. stellen fest, dass diese Studien hauptsächlich von universitären oder Internet-Service-Provider-afinen Institutionen durchgeführt werden und kein Bezug zu den Anbietern der sozialen Netzwerke selber besteht, weshalb sie den Studien eine beschränkte Aussagekraft zuschreiben (Jin et al., 2013, S.147). Die Datengrundlage der vorzustellenden Studien — und somit auch der Clickstreams — sind oft auf HTTP-Anfrage-Antwort-Paaren basierend, wodurch sich diese auch zur Netzwerkkanalyse eignen und aufzeigen, welche Funktion einer Social-Media-Plattform zu welchen Uhrzeiten den höchsten Netzwerkverkehr erzeugt etc., woraus auch infrastrukturelle Optimierungen abgeleitet werden können (Jin et al., 2013). Für den vorliegenden Kontext steht jedoch das generelle Nutzungsverhalten im Mittelpunkt, mit dem Ziel, die User-Experience der Nutzer verbessern zu können. Benevenuto et al. formulieren den Nutzen von auf Clickstream-Daten basierenden Studien wie folgt:

„Understanding how users behave when they connect to social networking sites creates opportunities for better interface design, richer studies of social interactions, and improved design of content distribution systems “ (Benevenuto et al., 2009, S. 49).

Benevenuto et al. nutzen die HTTP-Headerinformationen eines brasilianischen Netzwerkaggregators<sup>5</sup>, um detaillierte Clickstream-Daten von etwa 37 000 Nutzern auf den vier Social-Media-Anwendungen Orkut<sup>6</sup>, Myspace<sup>7</sup>, Hi5<sup>8</sup> und LinkedIn<sup>9</sup> zu analysieren. Haupterkenntnis ihrer Studie ist die Tatsache, dass bis zu 92% aller Nutzeraktivitäten auf Social-Media-Plattformen Browsing-Verhalten ist, das aus *latenten* bzw. *unsichtbaren* Aktivitäten besteht. Sie weisen darauf hin, dass bestimmtes Verhalten nicht mit *sichtbaren* Interaktionen allein — also Daten, die mit Hilfe von Crawling oder APIs gesammelt werden — analysiert werden kann und argumentieren, dass Studien, die nur mit *sichtbaren* bzw. API-Daten arbeiten, automatisch einer gewissen Verzerrung unterliegen (Benevenuto et al., 2009, S.58). In Bezug auf das Sessionverhalten der Social-Media-Nutzer zeigen Benevenuto et al., dass sowohl die Anzahl der Sessions als auch die Sessiondauer (Median: Orkut: 13,4 Min.; Hi5: 2,7 Min .; MySpace: 24 Sek.; LinkedIn: 3 Sek.) in ihrer Verteilung einen sehr starken *long tail* aufweisen (Benevenuto et al., 2009, S. 52). Ferner stellen sie fest, dass die Profile anderer Nutzer bei der Informationsgewinnung und -verbreitung in OSNs stark im Zentrum stehen. Mehr als 80% aller Inhalte, die von Nutzern betrachtet werden, stammen von direkten Freunden (Benevenuto et al., 2009, S. 61).

Im selben Jahr veröffentlichen Schneider und Kollegen eine ähnliche Studie, deren Datengrundlage und methodisches Vorgehen der Studie von Benevenuto et al. sehr ähnlich ist (Schneider et al., 2009). Auch hier bilden die HTTP-basierten Clickstream-Daten, welche via *reverse engineering*

<sup>5</sup>Ein Netzwerkaggregator ist eine Webiste, welche den Inhalt bzw. die Profile von mehreren Social-Media-Plattformen auf eine zentrale Umgebung zusammenfasst. Ein bekannter Vertreter ist *Hootesuite*.

<sup>6</sup>Schließung der Plattform September 2014

<sup>7</sup><https://myspace.com/>

<sup>8</sup><http://www.hi5.com/>

<sup>9</sup><https://www.linkedin.com/>

aus HTTP-Anfrage-Antwort-Paaren eines Internet-Service-Providers gewonnen werden, die Datengrundlage. Auch hier werden die Daten von mehreren Social-Media-Applikationen genutzt. Diese sind *Facebook*, *LinkedIn*, *Hi5* und *StudiVZ*<sup>10</sup>. Grundsätzlich stellen sie fest, dass bei unterschiedlichen OSNs unterschiedliche Funktionen im Vordergrund stehen. Beispielsweise ist die Interaktion mit Facebook-Funktionen heterogen im Vergleich zu Plattformen wie *StudiVZ* und *Hi5*. In Bezug auf die Anzahl an Sessions und deren Dauer stellen Schneider et al. fest, dass eine sehr starke Long-Tail-Verteilung vorliegt — einige dauern nur wenige Minuten, wenige dauern mehrere Stunden (Mittelwert: 40 Min.) — und bestätigen damit auch die Erkenntnisse von Benevenuto et al. (Schneider et al., 2009). Wie auch Benevenuto et al. heben Schneider und Kollegen die Bedeutung von Nutzerprofilen bei der OSN-Nutzung hervor. Sie zeigen, dass die Größe des Freundesnetzwerks kein guter Indikator für die Anzahl an besuchten Profilen in einer Session ist und die Anzahl an besuchten Profilen generell eine sehr schiefe Verteilung besitzt. (Schneider et al., 2009, S.36)

Wang et al. verwenden die Clickstream-Daten (etwa 7 Mio. Klicks über einen Zeitraum von 2 Monaten) von 16 000 Nutzern der chinesischen Social-Media-Plattform *Renren*<sup>11</sup>, um Klickmodelle zu erstellen, die zwischen falschen Identitäten bzw. böswillige Nutzeraccounts und normalen Nutzern unterscheiden (G. Wang et al., 2013). Sie zeigen, dass sowohl in Bezug auf verschiedene Sessioncharakteristika (Sessiondauer, Anzahl Sessions pro Nutzer pro Tag etc.), Nutzeraktivitäten sowie Übergangswahrscheinlichkeiten zwischen Klicks in bestimmten OSN-Kategorien (Markov-Ketten), normale Nutzer ein deutlich heterogenes Verhalten an den Tag legen, als sog. *Sybil*-Accounts dies tun. Böswillige Accounts haben sehr wenige Sessions mit wenigen Klicks, die überproportional häufig auf das Senden von Freundschaftsanfragen abzielen (G. Wang et al., 2013, S.243). In einem Klassifikationsexperiment, bei der eine *Support Vector Maschine* (SVM) als Klassifikator genutzt wird, können Wang und Kollegen unter Verwendung von 12 Features mit einer Genauigkeit von 98,9% normale Nutzer von böartigen Nutzern unterscheiden (G. Wang et al., 2013, S.245). Als wichtigstes Feature wird die Anzahl der Klicks für Freundschaftsanfragen identifiziert.

In einer erst kürzlich erschienenen Studie entwickeln Wang et al. eine Visualisierungsumgebung mit dem Ziel, das Nutzungsverhalten von OSN-Nutzerclustern zu visualisieren (G. Wang et al., 2016). Für einen Nutzertest mit 15 Probanden, in dem vor allem getestet werden soll, wie gut die Umgebung dazu dient, tatsächliches Nutzerverhalten aus diesen Clustern ablesen zu können, greifen Wang und Kollegen auf die Clickstream-Daten der beiden sozialen Netzwerke *Whisper*<sup>12</sup> und *Renren*<sup>13</sup> zurück. Anhand von drei Fallstudien wird das Verhalten von Nutzern auf *Whisper* und *Renren* auch konkret beschrieben. In einer Fallstudie zur Plattform *Whisper* untersuchen sie ein Cluster, in welchem die Clickstream-Aktionen der Nutzer einen überproportionalen Anteil an Events enthalten, die das Ziel verfolgen, andere Nutzer zu blocken. Sie vergleichen Nutzer innerhalb dieses Clusters und außerhalb des Clusters und stellen fest, dass der Grund für das häufige Blocken anderer Nutzer die Tatsache ist, dass diese Nutzer häufig *Whispers* mit erotischen oder pornografischen Inhalt veröffentlichen, mehr Nachrichten senden, mehr *likes* dafür erhalten und auch häufiger private Chatanfragen erhalten, woraus auch die verhältnismäßig hohe Blockquote resultiert.

Zusammenfassend ergeben sich aus den vorgestellten Studien folgende Hinweise für die eigene Logstudie:

- Es gibt nur wenige Studien, die mit Clickstream-Daten arbeiten, um das Verhalten von Nutzern auf Social-Media-Plattformen zu beschreiben. In vielen Studien erfolgt eine oberflächliche Beschreibung des Verhaltens. Wang und Kollegen zeigen jedoch, dass auf der Basis des

<sup>10</sup><http://www.studivz.net>

<sup>11</sup><http://www.renren.com/> Renren ist das chinesische Pendant zu Facebook.

<sup>12</sup><https://whisper.sh/>

<sup>13</sup>Dabei wird derselbe Datensatz verwendet, welcher bereits in der Arbeit von Wang et al. 2013 zur Kategorisierung von böswilligen Accounts genutzt wurde (G. Wang et al., 2013).



Nutzerverhaltens normale Nutzer von Nutzern mit bösartigen Absichten (Spam) unterschieden werden können.

- In Bezug auf das allgemeine Verhalten liefern diese Studien erste Hinweise darauf, wie lange Social-Media-Sitzungen sein können und welche Grenzwerte für die Identifikation von Sessions auf Social-Media-Plattformen sinnvoll sind. Der Wert von fünf Minuten als Sessiongrenze kann als erster Orientierungspunkt gelten.
- Die Studien zeigen, dass bis zu 92% des Nutzerverhaltens latentes Verhalten (z. B. Browsing von Unterseiten wie Nutzerprofilen) ist. Im Umkehrschluss bedeutet dies, dass von individuellen Personen nur 8% der Interaktion mit einer Social-Media-Plattform über APIs ausgelesen werden können. Um ein detailliertes, beinahe lückenloses Verhalten von Nutzern zu erfassen, muss auf andere Datenerhebungsmethoden als APIs zurückgegriffen werden.

## 6.3. Erhebung der Logdaten für die Twitter-Logstudie

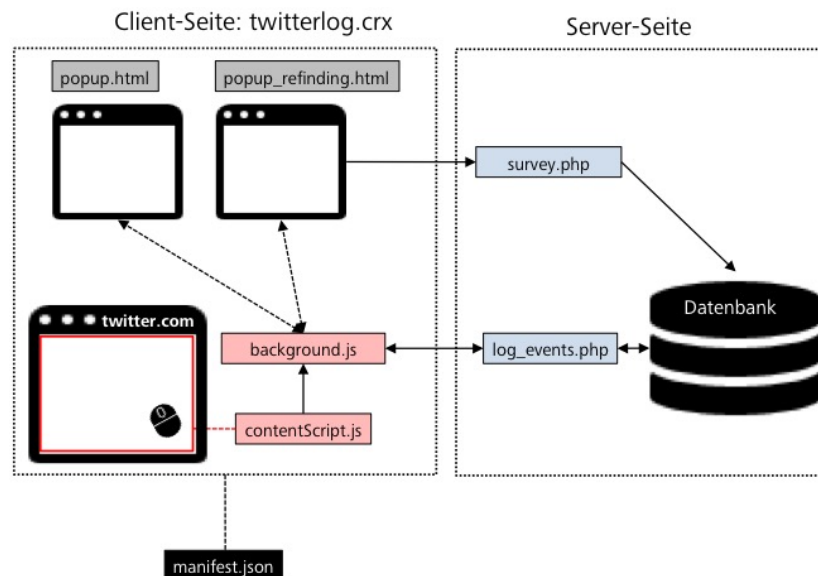
In Abschnitt 6.2.2 wurde gezeigt, dass Logdaten unterschiedlich detailliert sein können und ihre Erhebung auch unterschiedlich aufwändig sein kann. Um einen objektiven, realistischen Eindruck vom tatsächlichen Nutzerverhalten beim Wiederfinden von Tweets zu bekommen und eine Datenerhebung ohne Experimenteffekt zu gewährleisten, bietet sich die Erhebung von Clickstream-Daten an. Die in Abschnitt 6.2.5 vorgestellten Studien haben gezeigt, dass bis zu 92% aller Nutzerinteraktionen auf Social-Media-Plattformen latente Interaktionen — wie der Besuch von Profilen anderer Nutzer, oder der Besuch der eigenen Favoritenliste — sein können (Benevenuto et al., 2012; Jiang et al., 2010). Ziel der Clickstream-Datenerhebung in der Twitter-Logstudie ist es, ein komplettes, lebensnahes Interaktionsbild der Probanden zu erhalten und die Nachteile der API, die stets aktive Teilnahme, wie Favorisieren oder Weiterleiten von Tweets verlangt, zu umgehen. Somit wird auch passive bzw. latente Twitter-Nutzung — also das Lesen/Klicken von Tweets oder Navigation auf Unterseiten — aufgezeichnet. Für diesen Zweck wurde eine Erweiterung für den Chrome-Webbrowser als Erhebungsinstrument entwickelt. Die Entwicklung und Bestandteile der Erweiterung werden im Folgenden beschrieben.

### 6.3.1. Chrome Webbrowser-Erweiterung

Google Chrome-Nutzer besitzen die Möglichkeit, den Browser durch die Installation von Softwarekomponenten — sogenannten Erweiterungen (*extensions*) — um zusätzliche Funktionalität zu ergänzen. Chrome-Erweiterungen existieren für die unterschiedlichsten Nutzungskontexte beginnend bei Bedienhilfen (Inhalts-Blocker, Screen-Reader etc.), Entwicklerwerkzeugen (Echtzeit-CSS-Editoren), Wetter- oder Shopping-Applikationen bis hin zu Spielen. Häufig besitzen Erweiterungen kein oder nur ein sehr reduziertes UI, welches sich nahtlos als Icon in das UI des Browsers integriert. Erweiterungen können entwickelt werden, ohne in den nativen Code des Browsers eingreifen zu müssen. Sie basieren auf herkömmlichen Webtechnologien wie HTML5, JavaScript und CSS, wodurch sich die Möglichkeit ergibt, neben den APIs von Chrome auch diejenigen APIs zu nutzen, die der Browser für Webseiten zur Verfügung stellt z. B. XMLHttpRequests oder HTML5 APIs wie *canvas*, *geolocation* oder *local storage*. Sämtliche, für eine Erweiterung benötigten Dateien, werden in ein Paketformat mit der Dateiendung `.crx` gebündelt. Erweiterungen werden im Regelfall über den *Chrome web store*<sup>14</sup> zum Download angeboten. Ein Hosting auf eigenen Servern ist prinzipiell möglich, jedoch existieren aus Sicherheitsgründen gewisse Installationsbeschränkungen (Google,

<sup>14</sup><https://chrome.google.com/webstore/category/apps>

2016). Darüber hinaus erleichtert das Anbieten der Erweiterung über den *Chrome web store* das automatische Aktualisieren der Erweiterung (*auto updating*), weshalb dies den Regelfall darstellen sollte. Abbildung 6.2 zeigt die Architektur der Erweiterung, die zur Sammlung der Nutzerinteraktionsdaten mit der Twitter-Website genutzt wurde.



**Abbildung 6.2.:** Architektur der Chrome-Erweiterung zur Sammlung der Clickstream-Daten

Die Erweiterung verfügt über eine `manifest.json`-Konfigurationsdatei. In ihr sind wesentliche Eigenschaften der Erweiterung wie Sicherheitsaspekte (*content security policy*), verwendete APIs, `permissions` sowie Metadaten wie Name, Referenz auf das Logo, Angabe der Versionsnummer usw. gespeichert. Ferner ist angegeben, auf welchen Webseiten die Erweiterung aktiv sein darf (`content_scripts: „matches“`) und welche zusätzlichen Bibliotheken respektive JavaScript-Anwendungen aktiv sein dürfen.

Da es sich bei der Twitter-Log-Erweiterung nicht um eine Erweiterung handelt, die dem Browser Funktionalität hinzufügt, sondern den Zweck besitzt, ein Protokoll der Interaktionsevents zu generieren, verfügt sie nicht über *browser actions* oder *page actions*, welche im Normalfall durch spezifische UI Elemente realisiert sind und in der Toolbar des Browsers angezeigt wären. Somit existiert auch keine `background.html` respektive UI-Seite (*views*), welche für diese Zwecke notwendig wäre. Zentraler Bestandteil der Erweiterung ist das `contentScript.js`. *Content scripts* sind JavaScript-Dateien, die im Kontext der jeweiligen Webseite geladen werden und somit als direkter Bestandteil der Seite interpretiert werden können. Sie erlauben das Lesen und Verändern sämtlicher auf einer Webseite enthaltenen DOM-Elemente auch derjenigen, die erst durch Interaktion später zum DOM hinzugefügt werden. Dies wurde im Kontext der entwickelten Erweiterung ausgenutzt, indem zur Sammlung der Interaktionsevents sämtliche relevanten DOM-Elemente über CSS-Selektoren unter der Verwendung von JQuery mit einem sog. *listener* belegt wurden. Neben Klick- und Hover-Ereignissen wurden auch die Mausbewegungen (z. B. Anzahl der Bewegungen, Reichweite des Scrollens in Pixel) im Allgemeinen protokolliert. Darüber hinaus

wurden auch Texteingaben via Tastatur — für die Erstellung eines Tweets oder einer Suchanfrage bei der Twitter-Suche — aufgezeichnet. Ferner wurden wichtige Webseitenbereiche als sogenannte *areas of interest* (AOI) definiert. Ein Eindringen der Maus in einen AOI-Bereich wurde ebenfalls als AOI-Ereignis registriert.

Die Erfassung der Interaktionsereignisse geschah auf sehr detaillierte Art und Weise. Im Folgenden wird exemplarisch anhand der potentiellen Ereignisse, die bei der Interaktion mit einem Tweet entstehen können, zum einen der Detailreichtum der Logs veranschaulicht und zum anderen der Komplexitätsgrad, der durch die Beachtung dieser verschiedensten Ereignisse bei der Entwicklung bestand, verdeutlicht. Abbildung 6.3 zeigt einen Beispiel-Tweet und die Ereignisse, die durch einen Klick auf die jeweiligen UI-Elemente des Tweets ausgelöst werden.

Ereignis	Eindeutige CSS-Klasse(n)
1 RetweetedByClick	.js-retweet-text
2 TweetAuthorClick	a.account-group
3 TweetAuthorClick	a.account-group
4 TweetTimeClick	span._timestamp
5 TweetMentionClick	a.twitter-atreply
6 TweetClickonHashtag	a.twitter-hashtag
7 TweetClickonURL	span.js-display-url
8 TweetMediaClick	div.js-media-container
9 TweetReplyClick	.js-actionReply
10 TweetRetweetClick	.ProfileTweet-actionButton .js-actionRetweet
11 TweetUndoRetweet	.ProfileTweet-actionButtonUndo .js-actionRetweet
12 TweetFavorite	.ProfileTweet-actionButton .js-actionFavorite
13 TweetUnfavorite	.ProfileTweet-actionButtonUndo .js-actionFavorite
14 TweetShareViaEmailClick	li.share-via-email
15 TweetCopyLinkClick	li.js-actionCopyLinkToTweet
16 TweetEmbedTweetClick	li.embed-link
17 TweetMuteUser	li.mute-user-item
18 TweetBlockorReport	li.block-or-report-link
19 TweetBlockorReport	li.block-or-report-link
Weitere hier nicht abgebildete Interaktionsmöglichkeiten	
20 TranslateTweetClick	.js-translate-tweet
21 TweetGeoLocationClick	span.ProfileTweet-geo
22 ViewAllReplies	a.missing-tweets-link

**Abbildung 6.3.:** Überblick über die Interaktionsmöglichkeiten mit einem Tweet und die zugehörigen CSS-Klassen

Jeden Interaktionsschritt im Detail zu erfassen, bedarf einer tiefen Auseinandersetzung mit den jeweiligen CSS-Selektoren des Twitter-Web-Interfaces. Jedoch wird dadurch ermöglicht, dass bei der späteren Analyse eine differenzierte Betrachtung der Nutzerinteraktion durchgeführt werden kann und je nach Fragestellung bestimmte Interaktionen als Stellvertreter (Proxy) für Nutzerabsichten gewertet werden können. Ein Ereignis wie `TweetClickonHashtag` kann bei der Analyse von Suchanfragen als themenspezifische Suchanfrage gewertet werden, obwohl es sich eigentlich um einen Klick auf einen Link handelt. Im Kontext von Re-finding kann bei einem erneuten Klick auf einen Hashtag dieser Re-klick als Wiederfindensgrund interpretiert werden (siehe Abschnitt 6.4.4.7).

Die detaillierte Ereignisverarbeitung ist in Abbildung 6.4 dargestellt. Bei jedem Ereignis wird durch das `contentScript.js` ein JSON-Objekt erstellt, das die wesentlichen Informationen über das Ereignis wie Art des Ereignisses (`type`), Zeitstempel (`timestamp`) sowie Kontext (`context`), also die URL für die konkrete Unterseite, in der die Aktion stattfand, enthält. In Abhängigkeit von der Ereignisart können weitere Informationen als Objekte enthalten sein, wie im Beispiel des Er-



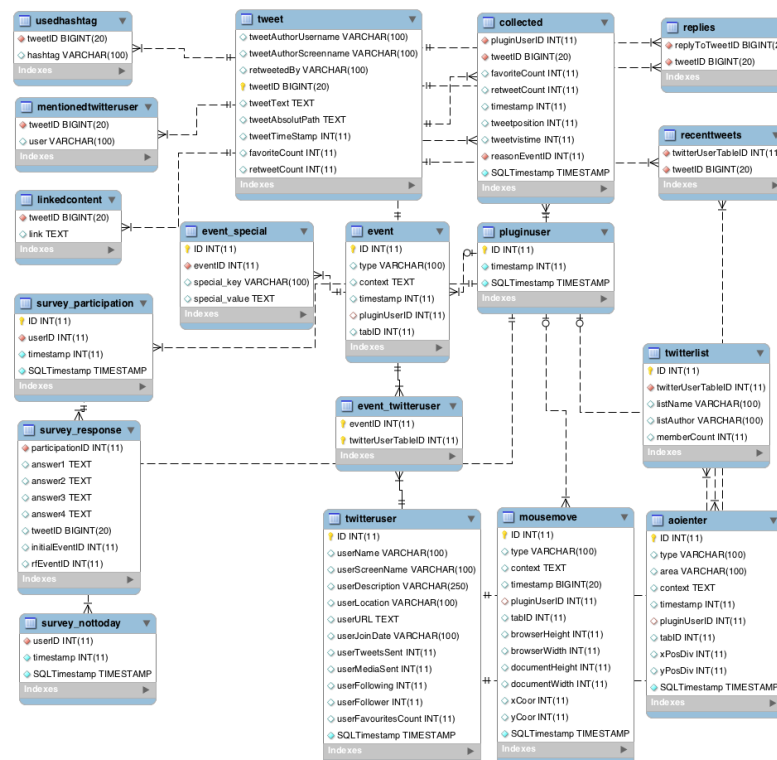


Abbildung 6.5.: Entity-Relationship-Modell der Datenbank zur Speicherung der Interaktionsevents

mehrfach in dieser Tabelle aufgeführt sind. Als *unique Tweets per User* ist schließlich die Zahl an Tweets zu verstehen, bei der derselbe Tweet von zwei unterschiedlichen Nutzern geklickt wurde. Diese Bezeichnungen werden erneut in Abschnitt 6.4.4 von Bedeutung sein. Zur Entwicklung der Erweiterung wurden zahlreiche Bibliotheken und Frameworks verwendet.

**JQuery:**<sup>15</sup> Im Rahmen der Erweiterung wurde sehr intensiv die Bibliothek *JQuery* eingesetzt, um ihre Möglichkeiten zur DOM-Traversierung und -Manipulation, dem Event-Handling für das Belegen der DOM-Elemente mit *listener* sowie der in ihr implementierten AJAX-Funktionalität zu nutzen.

**Mutation Summary:**<sup>16</sup> Die Bibliothek *Mutation Summary* baut auf der Browser-API *DOM Mutation Observer* auf und ermöglicht es, Veränderungen des Webseiten DOMs in Echtzeit schnell und speichereffizient zu beobachten und darauf zu reagieren. *Mutation Summary* liefert dabei eine Zusammenfassung des Zustands einer Webseite bis zum Zustand des nächsten *callbacks*. Diese Zusammenfassung kann ausgelesen und weiterverarbeitet werden. Eine *Query API* ermöglicht die Beobachtung von spezifischen Seitenelementen. Die Bibliothek wurde in der Twitter-Loggerweiterung für folgende Funktionalitäten verwendet:

- Erfassen der Tweets, die durch Scrollen auf der *Timeline* erscheinen
- Erfassen der Antworttweets (*@Replies*) auf einen Tweet, die nach Klick auf einen Tweet erscheinen
- Erfassen des Profilvorschau-Pop-ups (*profile summary*), die nach einem *@Mention*-Klick oder

<sup>15</sup><https://jquery.com/>

<sup>16</sup><https://github.com/rafaelw/mutation-summary>

einem Profilaufruf erscheinen

- Erfassen des Bildergalerie-Pop-ups (*media gallery*), das sämtliche Bilder zeigt, die Teil von gesendeten Tweets eines Nutzers sind
- Erfassen von Texteingaben wie Tweet-Antworten (@Replies), Kommentare bei Re-tweets oder Empfänger von E-Mails und Direktnachrichten
- Erfassen von iFrame-Inhalten wie Zusammenfassungen von Website-Nachrichten als Tweet-Anhang

Durch Interaktion neu zur Website hinzukommende DOM-Elemente mussten ebenfalls neu mit *listener* versehen werden, um Interaktionen mit diesen neuen Elementen erfassen zu können.

**Mousetrap:**<sup>17</sup> Twitter ermöglicht auch die Nutzung via Tastaturkürzeln. Um sicherzustellen, dass auch eine mausfreie Interaktion über diese Kürzel registriert werden konnte, wurde die Bibliothek *Mousetrap* verwendet. *Mousetrap* ermöglicht eine einfache Verarbeitung von Tastaturereignissen.

**VisSense.js:**<sup>18</sup> *VisSense.js* ist eine Bibliothek, die es erlaubt, die Sichtbarkeit von DOM-Elementen zu überprüfen. Sie bietet eine sog. Convenience-Klasse, welche mit Hilfe von verschiedenen Methoden (z. B. `isVisible()`, `isHidden()`) die Sichtbarkeit eines DOM-Elements auf dem Bildschirm des Nutzers überprüft. Sie wurde dazu genutzt, um die Dauer der Sichtbarkeit eines Tweets bis zu Interaktion zu messen.

**Bootstrap:**<sup>19</sup> Die für das *Experience Sampling* genutzten HTML-Dateien wurden mit Hilfe des Frontend-Frameworks *Bootstrap* gestaltet.

**Propel:**<sup>20</sup> Serverseitig wurde die PHP-Bibliothek *Propel* genutzt. *Propel* ist eine *PHP Data Objects* (PDO) basierte Object-relational-mapping-Bibliothek (ORM), die eine objektrelationale Abbildung zwischen den durch die Erweiterung erzeugten Ereignisobjekten und der relationalen Struktur der Datenbank ermöglicht. Unter Verwendung von Objekten und einer API zum Speichern und Lesen von Daten werden die Ereignisse in der MySQL-Datenbank gespeichert. Sowohl das Erzeugen des Datenbankschemas, das im Falle von *Propel* über XML gelöst werden kann, sowie das Speichern und Lesen der Daten wurde durch *Propel* erheblich vereinfacht und ist im Vergleich zu normalen SQL-Anfragen robuster und weniger fehleranfällig.

### 6.3.2. Experience Sampling

Ein entscheidender Nachteil von lebensnahen, quantitativ orientierten Logstudien ist der Mangel an qualitativen Daten, die über Nutzerabsichten informieren, Kontextinformationen liefern und somit bei der Interpretation der quantitativen Daten helfen. Die Ergänzung von Logdatenstudien um zusätzliche Erhebungsmethoden, wie klassische Fragebögen, ist ein weit verbreitetes und empfohlenes Vorgehen (Griffiths et al., 2002; Lazar et al., 2010). Für den vorliegenden Zweck wurde die Twitter-Logstudie um die qualitative Erhebungsmethodik des *Experience Sampling* erweitert.

#### 6.3.2.1. Die Experience-Sampling-Methode

Die Experience-Sampling-Methode (ESM)<sup>21</sup> ist eine Methode zur Sammlung von Handlungsbeschreibungen, Gedanken und Gefühlen in alltäglichen Situationen in Form von Berichten, soge-

<sup>17</sup><https://craig.is/killing/mice>

<sup>18</sup><https://github.com/vissense/vissense>

<sup>19</sup><http://getbootstrap.com/>

<sup>20</sup><http://propelorm.org/>

<sup>21</sup>Der deutsche Begriff „Erlebens-Stichproben-Methode“ findet nur selten Anwendung.

nannten *self-reports*, die oft die Form eines kurzen Fragebogens besitzen. Die Sammlung von Berichten geschieht meist über einen längeren Zeitraum hinweg, weshalb Bolger und Laurenceau die ESM in die Kategorie der *intensive longitudinal methods* einordnen (Bolger & Laurenceau, 2013, S.11). Die ESM hat ihre Ursprünge in verschiedenen psychologischen Studien, respektive Methoden (Csíkszentmihályi & Larson, 2014, S.21). Erste ESM-artige Studien — wie die Untersuchung *How Working Men Spend Their Spare Time* von Bevens, welche die Freizeitgestaltung von Männern thematisiert — fokussieren auf sozialwissenschaftliche/soziologische Fragestellungen und lassen sich auf den Anfang des 20. Jhd. datieren (Bevens, 1913). Larson und Csíkszentmihályi sehen im *Experience Sampling* eine methodische Weiterentwicklung von Beobachtungsmethoden und Tagebuchstudien, die — im Falle von Beobachtungsstudien — nur dazu fähig sind, öffentliches Verhalten zu observieren, wodurch privates oder sehr persönliches Verhalten verloren geht und — im Falle von Tagebuchstudien — den Fokus zu sehr auf reines Verhalten legen und die gedankliche und emotionale Ebene, die Personen in unterschiedlichen Lebenssituationen charakterisieren, vernachlässigen (Csíkszentmihályi & Larson, 2014, S.22). Für Bolger und Laurenceau spielt auch der natürliche, spontane Kontext, der mit der ESM berücksichtigt werden kann, eine Rolle, da der Effekt des schlechten oder falschen Erinnerns von Gedanken und Gefühlen (*recall bias*) minimiert wird. Die Rekonstruktion von Ereignissen seitens des Probanden entfällt und führt zu einer höheren Objektivität der Daten (Bolger & Laurenceau, 2013, S.5) (Csíkszentmihályi & Larson, 2014, S.23).

Aus der Perspektive des experimentellen Designs können vier Arten von ESM-Studien unterschieden werden: intervallbasierte, signalbasierte, ereignisbasierte und — aus neuen technologischen Bedingungen hervorgegangen — endgerätebasierte Studien (Bolger & Laurenceau, 2013, S.14). Bolger und Laurenceau fassen intervallbasierte und signalbasierte Studien zu zeitbasierten Studien zusammen und stellen diesen die ereignisbasierten Studien gegenüber (Bolger & Laurenceau, 2013, S.14).

#### **Zeitbasiertes Experimentdesign**

Inintervallbasierte Studien, die die ESM verwenden, sind so angelegt, dass zu ein bis zwei festen Zeitpunkten am Tag der Proband dazu angehalten ist, einen fragebogenartigen Bericht abzugeben, der Informationen über Handlungen und Gefühle, die sich seit dem letzten Bericht ereignet haben, beinhaltet (Bolger & Laurenceau, 2013, S.15). Die Zeitpunkte werden meist zu Beginn und am Ende des Tages gesetzt, um Veränderungen über den Tagesverlauf beobachten zu können. Dieses Studiendesign wird auch *daily process designs* bezeichnet und laut Bolger und Laurenceau in unterschiedlichen thematischen Kontexten wie Persönlichkeitsentwicklung, Gesundheits- und Emotionsstudien, sowie in der Psychotherapie eingesetzt (Bolger & Laurenceau, 2013, S.15). Da die Berichterstattung durch den Probanden zu festen, wiederkehrenden Zeitpunkten und nur ein- bis zweimal täglich erfolgt, wirkt diese Studienart eher unaufdringlich und wird von Probanden gut akzeptiert. Jedoch wird durch die eher weiter auseinanderliegenden Berichterstattungszeitpunkte mehr Retrospektive und Erinnerungsvermögen vom Probanden verlangt (Bolger & Laurenceau, 2013, S.16). Im Gegensatz zu intervallbasierten Studien erfolgt der Bericht bei signalbasierten Studien nach dem Erhalt eines Signals, z. B. durch einen Pager, welches den Nutzer zum Erstellen eines Berichts auffordert. Die Signale sollten dabei spontan, ohne Vorwarnung versendet werden und müssen repräsentativ sein, also zu unterschiedlichen Zeitpunkten mehrmals am Tag erfolgen (Csíkszentmihályi & Larson, 2014, S.23ff.). Das signalbasierte Studiendesign wurde vor allem von Csíkszentmihályi und Larson entwickelt und genutzt, weshalb deren Studiendesignvarianten prototypisch für diese ESM-Variante gelten können (Bolger & Laurenceau, 2013, S.17). Die In-situ-Berichterstattung — als Folge randomisierter Aufforderungen mehrmals über den Tag verteilt — liefert sehr gute subjektive Einschätzungen über Verhalten, Erfahrung und Gefühle, kann jedoch aufgrund ihrer Häufigkeit als störend empfunden werden (Bolger & Laurenceau, 2013, S.17).

### Ereignisbasiertes Experimentdesign

Beim ereignisbasierten Experimentdesign werden Nutzer dazu aufgefordert, einen Bericht auszufüllen, sobald ein vordefiniertes Ereignis stattgefunden hat. Voraussetzung dafür ist, dass diese Ereignisse eindeutig vordefiniert sind und Probanden auch in der Lage sind, dieses Ereignis zu erkennen und sich schlussendlich auch zum Ausfüllen eines Berichts verpflichtet fühlen (Bolger & Laurenceau, 2013, S.18). Beispiel für klassische ereignisbasierte ESM-Studien sind die Rochester-interaction-record-Studien, in denen Probanden dazu angehalten waren, einen strukturierten Fragebogen zu sozialen Interaktionen auszufüllen. Sobald sie im Alltag eine soziale Interaktion durchgeführt haben, die länger als 10 Minuten gedauert hat, mussten Probanden einen Fragebogen beantworten. Anders als bei den vorherigen Studiendesigns kann die Einhaltung der Studienvorgaben nicht überprüft werden. Neue Endgeräte wie Smartphones tragen dazu bei, ereignisbasierte ESM-Studien leichter durchzuführen und zu kontrollieren.

Über verschiedene Endgeräte und deren Sensoren können Informationen bzw. Daten in unterschiedlichen Modalitäten in das *Experience Sampling* integriert werden. Vorstellbar sind: selbstgegebene Informationen durch Fragebögen, physiologische Daten (Koppelungen mit weiteren Endgeräten wie Aktivitäts-Tracker (*Fitbit*)), Erfüllung von Aufgaben, Umwelteinflüsse (Messung der Umgebungsgeräusche oder Temperatur) oder räumliche Informationen (GPS) sowie weitere Sensordaten des Smartphones (Bolger & Laurenceau, 2013, S.19). Dies führt zu einem kontextsensitiven Stichprobenverfahren, da der Nutzer nicht nur vom dem Druck befreit ist, Ereignisse zu erkennen, sondern diese durch das Endgerät zum adäquaten Zeitpunkt automatisch erkannt werden können.

Die ESM hat sich bei der Untersuchung des Informationssuchverhaltens etabliert. Case und Given führen die ESM als Standardverfahren zur Untersuchung von Informationssuchverhalten an und nennen die Studien von Spencer (Spencer, 1971) und Shim (Shim, 2003) als Beispiele (Case, 2007, S.222-249). Ein aktuelles Studienbeispiel mit ESM-Einsatz ist die Arbeit von Church, Cherubini und Oliver, die eine großangelegte ESM-Studie durchführen, um die täglichen Informationsbedürfnisse von Smartphone-Nutzern *in situ* zu erheben (Church et al., 2014). Diese Perspektive des *Experience Sampling* steht in sehr engem Zusammenhang mit dem sogenannten *lifelogging*, also der Nutzung von digitalen Geräten zur Aufzeichnung aller Facetten des Lebens über einen längeren Zeitraum hinweg in verschiedenen Detailgraden (Selke, 2016). Im Kontext des IR werden in den letzten Jahren vermehrt Life-Logging-Studien durchgeführt. Für die NTCIR wurde ein *pilot task* zur Life-Log-Analyse eingerichtet, welcher zwei Aufgaben/Tasks umfasst, nämlich der *lifelog semantic access task (LSAT)*, in denen speziell definierte Momente — z. B. wenn sich die Probanden in der Küche aufhalten — aus den Daten extrahiert werden müssen und der *lifelog insight task (LIT)*, bei dem Erkenntnisse über die Daten gewonnen werden sollen. Darüber hinaus wurde die erste Testkollektion für Life-Log-basierte Untersuchungen von der NTCIR auf der SIGIR 2016 veröffentlicht (Gurrin et al., 2016). Auch in benachbarten Feldern wie der *Human Computer Interaction* ist die ESM fester Bestandteil des qualitativen Methodeninventars. Hier wird ebenfalls auf den Vorteil hingewiesen, den mobile Geräte für die kontextuelle ES-Erhebung besitzen (Consolvo & Walker, 2003; Froehlich et al., 2007). Des Weiteren werden auch Verbesserungsvorschläge für die Methodik per se gegeben (Cherubini & Oliver, 2009).

#### 6.3.2.2. *Experience Sampling* in der Twitter-Loggerweiterung

Das *Experience Sampling* stellt einen weiteren Bestandteil der Erweiterung dar. Im Falle der Twitter-Loggerweiterung handelt es sich um ereignisbasiertes ESM-Design, wie im vorherigen Abschnitt beschrieben. Ein erneuter Klick auf einen Tweet, kann als Intention des Wiederfindens interpretiert werden und bildet das Ereignis, das ein Experience-Sampling-Pop-up evoziert. Unter Erfül-



**(a) Why did you visit that Tweet again?**

Please give a short description of your Twitter behaviour right now...

**Engadget** @engadget  
WSJ: Retina Display MacBook Air is coming soon  
engr.co/1M8arqx pic.twitter.com/YWJ9FrxV0  
10:17 AM - 5 Mar 2015  
30 retweets 23 favorites

1. Was it your intention to find this Tweet again?  
☐ Yes! This was re-finding!  
☐ No! This wasn't re-finding!  
☐ I'm not sure, this must be an error

2. If it was re-finding, how did you manage to get back to that Tweet?  
 Example: 'I remember Paul tweeting it so I looked at his Tweets...'

3. If it was re-finding, what was your motivation to get back to that Tweet?  
 Example: 'There was a cool link in it I wanted to visit again...'

**Submit** **Ask me another time!**

**(b) Questions about your re-finding experience on Twitter**

1. Have you recently had the need to find a certain Tweet again?  
☐ Yes! I currently had the need to re-find a Tweet.  
☐ No! I haven't had the need to re-find a Tweet.

2. If you had the need for re-finding, was your search successful?  
 Example: 'Yes it was easy to find...'

3. If your search was successful, how did you manage to get back to that Tweet?  
 Example: 'I remember Paul tweeting it so I looked at his Tweets...'

4. What was your motivation to get back to that Tweet?  
 Example: 'There was a great picture in it...'

**Submit** **Ask me another time!**

**Abbildung 6.6.:** Die Pop-ups der beiden Experience-Sampling-Varianten in Gegenüberstellung.

lung gewisser Heuristiken führte folglich ein erneuter Klick auf einen Tweet zu einem Experience-Sampling-Pop-up. Folgende Heuristiken kommen dabei zum Einsatz:

- Erste Interaktion mit Tweet muss 30 Minuten zurückliegen
- Nur ein Pop-up innerhalb von zwölf Stunden
- Nur fünf Pop-ups innerhalb von sieben Tagen
- Falls sich ein Nutzer gegen die Teilnahme an einem *Experience Sampling* ausspricht, erfolgt für 24 Stunden kein Pop-up.

Für den Fall, dass durch einen Nutzer sieben Tage lang kein erneuter Klick auf einen Tweet durchgeführt wurde, wurde ein alternatives *Experience Sampling* vorgelegt, das den Nutzer nach seinem allgemeinen Wiederfindenserfahrungen auf Twitter befragte. Technisch wird das *Experience Sampling* über die Datei `survey.php` geregelt, die anhand der letzten Datenbankeinträge auf Basis der Heuristiken überprüft, ob die jeweilige Versuchsperson die Heuristiken erfüllt, um an einem *Experience Sampling* teilzunehmen und die entsprechende Information an das `background.js` zurückführt, welche das entsprechende Pop-up aktiviert (vgl. Abbildung 6.2). Abbildung 6.6 zeigt die beiden Experience-Sampling-Optionen im Vergleich. Es ist zu erkennen, dass bei der Option mit explizitem Re-klick auch der erneut geklickte Tweet gezeigt wird. Prinzipiell waren die Fragen aber ähnlich. So wurde stets danach gefragt *Wie* Nutzer wiederfinden und *Warum* sie diesen Tweet bzw. allgemein Tweets wiederfinden.

### 6.3.3. Teilnehmerge Gewinnung und Rekrutierungsstrategien

Am 20. März 2015 wurde die Twitter-Loggerweiterung mit Hilfe des Entwickler-Dashboard auf dem *Chrome web store* veröffentlicht.<sup>22</sup> Dabei wurden die Sichtbarkeitsoptionen so gewählt, dass die Erweiterung nur für Personen erreichbar war, die über den direkten Link zur Erweiterung im *Webstore* verfügten. Auf eine uneingeschränkte Veröffentlichung wurde verzichtet.

<sup>22</sup><https://goo.gl/HVXMux>

Im Anschluss an die Veröffentlichung wurden folgende Rekrutierungsstrategien zur Teilnehmergewinnung verfolgt:

- Hinweis auf Logstudie über Mitteilungen in Social Media (Twitter und Facebook)
- Hinweis auf Logstudie in diversen Foren (u.a. Twitter-Entwickler-Forum, Chrome-Entwickler-Forum)
- Hinweis auf Logstudie durch öffentliche Aushänge und Verbreitung von Flyern
- Hinweis auf Logstudie in Lehrveranstaltungen der Informationswissenschaft und Medieninformatik
- Hinweis auf Logstudie über E-Mail-Verteiler

Zur extrinsischen Motivation wurde der Gewinn mehrerer Amazon-Gutscheine durch Teilnahme an einer Verlosung am Ende der Studie in Aussicht gestellt. Vor dem Download der Erweiterung wurden Nutzer über eine Teilnehmererklärung darüber aufgeklärt, welche Art der Daten aufgezeichnet werden und wie diese verwendet werden. Ferner stimmten die Nutzer einer Zuordnung zum jeweiligen Twitter-Account-Informationen sowie einer Einholung von weiteren Nutzerdaten via Twitter API zu. Insgesamt konnten 44 Nutzer rekrutiert werden, welche ein *convenience sample* bilden. Eine detaillierte Beschreibung der Studienteilnehmer erfolgt in Abschnitt 6.4.2

## 6.4. Ergebnisse der Logdatenauswertung

Die folgenden Abschnitte fokussieren auf die Auswertung der erhobenen Daten. Abschnitt 6.4.1 thematisiert die Vorarbeiten, die nötig waren, um die Logdaten für weitere Analysen aufzubereiten. Neben allgemeinem Twitter-Verhalten in Abschnitt 6.4.3 wird in Abschnitt 6.4.4 der Fokus auf Wiederfindensverhalten gelegt.

### 6.4.1. Vorarbeiten bei der Datenauswertung und -bereinigung

Vor der inhaltlichen Auswertung der Logdaten in Bezug auf die formulierten Leitfragen mussten einige Skripte (`cleanDupsUsers.R`, `cleanTypeContext.R`) zur Bereinigung der Logs entwickelt werden. Eine Diskrepanz bestand zwischen der synchronen Interaktion der Nutzer mit der Website und der asynchronen Kommunikation zwischen Erweiterung und Server. Die Diskrepanz zwischen Interaktion und Kommunikation mit dem Server führte in seltenen Fällen dazu, dass einzelne Events wiederholt hintereinander in die Datenbank eingetragen wurden, oder Inhalte nicht übertragen werden konnten. Diese Einträge wurden identifiziert und bereinigt. Darüber hinaus konnten ebenfalls Fehler bei der Zuordnung von Eventarten und Kontext (URL) festgestellt werden. Beispielsweise wurden manche Ereignisse fälschlicherweise als `userPageVisit` gewertet, obwohl die Interaktion auf der *Timeline* (Kontext: `https://www.twitter.com/`) stattfand. Ein größeres Problem stellt die automatische Synchronisierung von Daten wie Lesezeichen, Browserverlauf, Passwörtern und weiteren Einstellungen dar, die vom Chrome Webbrowser standardmäßig (Default-Einstellung) vorgenommen wird. Ist diese nicht deaktiviert, führt sie dazu, dass Erweiterungen über mehrere Geräte hinweg synchronisiert werden, respektive eine erneute Installation der Erweiterung auf einem anderen Gerät vorgenommen wird. Nutzer mit bestehenden `pluginUserIDs` erhalten durch die erneute Installation folglich eine weitere Nutzer-ID. Die Logdaten wurden hinsichtlich dieses Problems intellektuell bereinigt, indem anhand von bestimmten Events (z. B. Senden eines Tweets) ein Mapping zwischen verschiedenen `pluginUserIDs` durchgeführt wurde. Letztendlich ist aber nicht auszuschließen — dies zeigt sich auch in der Charakteri-

sierung der Studienteilnehmer in Abschnitt 6.4.2 — dass einige Studienteilnehmer durch mehrere Nutzer-IDs repräsentiert sind und die Zahl an einzelnen Versuchsteilnehmern geringer ist, als die absolute Zahl der IDs. Ein Umstand, der auch anhand der geringen Eventzahl für manche Teilnehmer zu erklären ist. In einem finalen Schritt wurden Nutzer-IDs mit Twitter-Account-Namen abgeglichen. Zusammenfassend wurden folgende Bereinigungen/Vorarbeiten vorgenommen:

- Löschung redundant-registrierter Events
- Behebung falscher Event-Kontext/URL-Zuweisungen
- Mapping zwischen Nutzer-IDs
- Mapping zwischen ID (`pluginUserID`) und Twitter-Namen

#### 6.4.1.1. Eventkategorisierung

Einzelne Ereignisse/Events wurden — analog zur Verfahrensweise in anderen Social-Media-Logstudien — nach Art des Ereignisses bzw. dem Kontext (URL) in dem das Ereignis stattfand, zu Ereigniskategorien zugeordnet. Beispielsweise wurden Ereignisse auf der eigenen *Timeline* (Kontext: <https://twitter.com>) als *TIMELINE*-Events kategorisiert. Hierzu gibt es einige Ausnahmen, die Vorrang vor kontextspezifischen Ereignissen besitzen: Ereignisse bestimmter Art, wie z. B. die Eingabe einer Suchanfrage, wurden unabhängig vom jeweiligen Kontext als zur Eventkategorie *SEARCH* zugehörig klassifiziert. Folgende Ereigniskategorien wurden verwendet:

- **TIMELINE:** Ereignisse im Kontext der eigenen *Timeline* (URL: <https://twitter.com/>) wurden, wie bereits erwähnt, als *Timeline*-Events gewertet.
- **NOTIFICATIONS:** Ereignisse im Kontext <https://twitter.com/notifications> sind Ereignisse, die zur Kategorie *NOTIFICATIONS* gerechnet werden.
- **OWNPROFILEINT:** Ereignisse auf dem eigenen Profil (URL: [https://twitter.com/\[OwnUserName\]](https://twitter.com/[OwnUserName])) werden zur Kategorie *OWNPROFILEINT* gezählt.
- **USERPROFILEINT:** Interaktionen auf den Profilseiten von anderen Nutzern (URL: [https://twitter.com/\[UserName\]](https://twitter.com/[UserName])) führen zu *USERPROFILEINT*-Events.
- **FFMRELATIONS:** FFM steht für *Follower-Following-Mention* und bezeichnet Interaktionen, die mit Personen-Interaktion in Zusammenhang stehen, ergo ein Klick auf eine *@Mention*, das Folgen und Entfolgen von neuen Profilen oder Ereignisse auf der Follower- oder Freunde-Seite im Kontext <https://twitter.com/followers> respektive im Kontext <https://twitter.com/following>.
- **SEARCH:** Neben der expliziten Eingabe einer Suchanfrage führen ein Klick auf einen Hashtag, ein Klick auf einen Link des Trendbereichs sowie weitere Interaktionen auf der Ergebnisseite (*SERP*-Seite) zu *SEARCH*-Ereignissen.
- **FAVOURITES:** Ereignisse auf der eigenen Favoritenliste (URL: <https://twitter.com/favorites>) wurden als *FAVOURITES*-Ereignisse kategorisiert. Events auf den Favoritenlisten von anderen Nutzern (URL: [https://twitter.com/\[Username\]/favorites](https://twitter.com/[Username]/favorites)) gehören zur Kategorie *USERPROFILEINT*.
- **SINGLETWEET:** Ereignisse, die im Kontext einzelner Tweets stattfanden (zu erkennen an URLs mit *TweetIDs*), wurden zur Eventkategorie *SINGLETWEET* gerechnet.
- **TWEETING:** Das Absenden eines Tweets resultierte in einem *TWEETING*-Event.

- **DIRECTMESSAGE:** Ein Klick auf das *Messages*-Icon führte zu einem Event der Kategorie DIRECTMESSAGES.
- **LISTS:** Ereignisse, die in Kontexten stattgefunden haben, welche in ihrer URL das Keyword *lists* enthalten wie z.B: `https://twitter.com/[username]/lists` oder `https://twitter.com/[username]/lists/[list-name]` wurden als LISTS-Events kategorisiert.
- **OTHER:** Ereignisse, die sich im Kontext einer Twitter-Subdomain (URL: `http://api.twitter.com` oder `dev.twitter.com`) ereigneten, wurden als Kategorie OTHER klassifiziert.

#### 6.4.1.2. Einteilung der Daten in Sessions

Logdaten werden üblicherweise in Nutzungsabschnitte oder Sitzungen, sog. Sessions, eingeteilt. Die Twitter-Website bietet eine Art *single sign on experience* (SSO), was zur Konsequenz hat, dass Nutzer, die sich einmal mit Nutzernamen und Passwort angemeldet haben auch angemeldet bleiben, falls sie keinen expliziten Log-out vornehmen. Folglich existieren beinahe keine Log-in Ereignisse, weshalb es nicht möglich ist, einzelne Sitzungen anhand dieser zu erkennen bzw. die Daten anhand dieser Ereignisse in Sessions einzuteilen. Sessiongrenzen zu bestimmen ist somit ein wesentlicher erster Schritt bei der Vorverarbeitung der Daten. Für die Einteilung der Twitter-Logdaten in einzelne Sessions/Sitzungen wurde eine Aktivitätspause von fünf Minuten als Sessiongrenze gewählt. Liegen zwei Interaktionen mehr als fünf Minuten auseinander, so befinden sie sich gemäß dieser Definition in unterschiedlichen Sessions. Die Wahl von fünf Minuten ist dabei einerseits auf Basis relevanter Literatur begründet, andererseits empirisch fundiert:

##### Begründung aus der Literatur

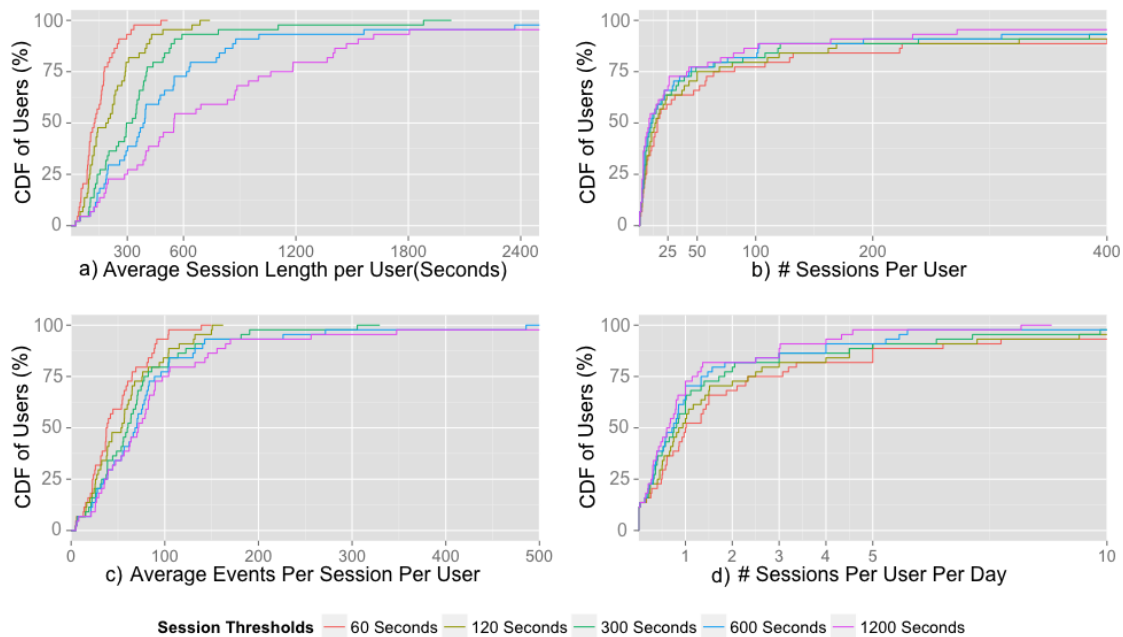
Wie bereits erwähnt, ist die dreißigminütige Inaktivität in der Web-Logdatenanalyse eine häufig genutzte Zeitspanne, um Sessions voneinander abzugrenzen (Jansen, 2006). Aber auch andere zeitliche Grenzen, wie eine fünfminütige Pause zwischen zwei Suchanfragen, wird im Webkontext als Ende respektive Beginn einer neuer Session gewertet (Silverstein et al., 1999). In anderen Kontexten werden ebenfalls zeitbezogene Inaktivitätsgrenzen unterschiedlicher Dauer genutzt. So werden bei Studien zu Desktopsuchmaschinen (Cutrell et al., 2006) oder E-Mail-Clients (Elsweiler, Harvey & Hacker, 2011) häufig fünf Minuten Inaktivität als Zeichen für das Ende bzw. den Beginn einer neuen Session gewertet. Auch bei Studien zu Social-Media-Logdaten bzw. Clickstream-Daten wird die Fünf-Minuten-Grenze zur Separierung der Daten in Sessions eingesetzt (Schneider et al., 2009).

##### Begründung aus den Daten

Um für die Analyse der Twitter-Logdaten einen passenden Session-Grenzwert zu finden, wurden unterschiedliche zeitbasierte Sessiongrenzen (eine Minute, zwei Minuten, fünf Minuten, zehn Minuten und zwanzig Minuten) verwendet und in Bezug auf verschiedene sessionspezifische Features wie *Anzahl der Sessions pro Nutzer* oder *Durchschnittliche Sessionlänge pro Nutzer* als empirische/-kumulative Verteilungsfunktionen <sup>23</sup> visualisiert (Abbildung 6.7). Je länger die Inaktivitätsgrenze ausfällt (z. B. eine Minute vs. zwei Minuten), desto länger wird die durchschnittliche Dauer einer Sitzung. Je kürzer die Inaktivitätsgrenze gewählt wird, desto mehr Einzelsitzungen werden erzeugt. Als Problem hieraus ergibt sich, dass bei der Wahl einer zu kurzen Sessiongrenze mögliche noch andauernde Sessions abgebrochen werden und als neue Sessions gewertet werden bzw. besteht die Gefahr, dass bei der Wahl einer zu langen Sessiongrenze mehrere einzelne Sessions zu einer Sitzung zusammengefasst werden, obwohl diese für sich genommen eigene Sitzungen

<sup>23</sup>Für jeden Punkt auf der X-Achse wird der Prozentsatz der Daten gezeigt, der diesem Wert entspricht. Somit bildet es eine Mischung aus Boxplot und Histogramm.

darstellen. In Abbildung 6.7 a) ist zu erkennen, dass die Wahl der Inaktivitätsgrenze zwischen



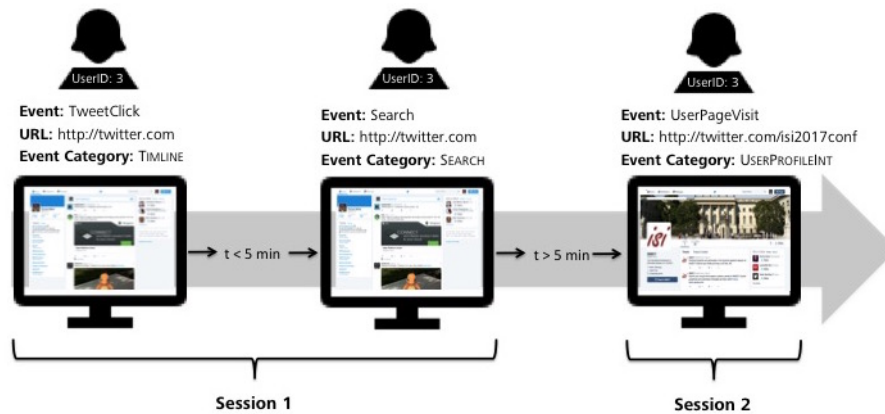
**Abbildung 6.7.:** Sessionspezifische Nutzerfeatures mit unterschiedlichen Inaktivitätsgrenzen im Vergleich

einer und fünf Minuten eine vergleichbare durchschnittliche Sessionlänge für die Nutzer ergibt, wohingegen bei der Wahl der Interaktionspause von 10 Minuten oder länger eine deutlich längere durchschnittliche Sessionlänge resultiert. Betrachtet man Abbildung 6.7 b) wird deutlich, dass für eine bis zwei Minuten als Interaktionspause die Anzahl der Einzelsessions pro Nutzer höher ausfällt, wobei sich bei der Wahl von fünf Minuten (grüne Linie) der Wert der Einzelsessions an die Zahl der langen Sessionintervalle annähert. Eine Interaktionspause von fünf Minuten kann so interpretiert werden, dass ihre Wahl nicht unnötig viele Einzelsessions (Nähe zu zehn und zwanzig Minuten in Abbildung 6.7 b)) produziert, wobei diese auch nicht zu lange dauern (Nähe zu ein und zwei Minuten in 6.7 a)). Die Abbildungen 6.7 c) und d) lassen ebenso keine Auffälligkeiten erkennen, die gegen fünf Minuten als Sessiongrenze sprechen.

Einen weiteren datengetriebenen Ansatz für die Wahl der Sessiongrenze bietet die Herangehensweise vermeintlich zeitintensive Interaktionen, wie das Lesen eines Artikels oder das Betrachten eines Videos, welche in einem Tweet verlinkt sind, als maximale Pause zwischen Folgeinteraktionen zu interpretieren. Eine Sessiongrenze dürfte diesen Wert nicht unterschreiten. Betrachtet man die Zeitspanne zwischen diesen Ereignissen und einer Folgeaktion ergibt sich ein Mittelwert von 27,9 Sekunden (*Median* = 5 Sek.). Als Orientierungspunkt für die Wahl der Sessiongrenze dient das Maximum, welches bei 294 Sekunden (4,9 Min.) und somit nahe an der Fünf-Minuten-Grenze liegt.

Fasst man die genannten Argumente aus Literatur und empirischer Analyse zusammen, ist die Wahl einer fünfminütigen Inaktivität als Sessiongrenze wohl fundiert. Wendet man diese Sessiongrenze auf die Daten an, ergeben sich insgesamt 4573 Sessions. Die Sessioncharakteristika pro

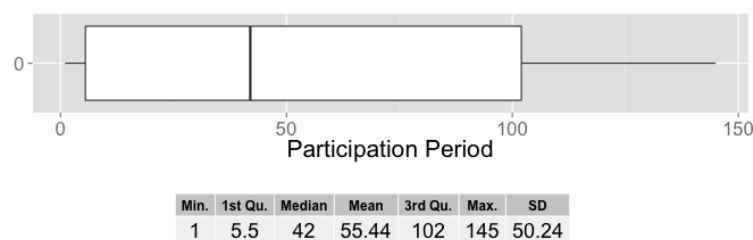
Nutzer werden in Abschnitt 6.4.3 dargestellt. Abbildung 6.8 fasst noch einmal die wesentlichsten Schritte der Vorverarbeitung — Eventkategorisierung, Sesssioneinteilung — visuell zusammen. Im Folgenden werden die Teilnehmer der Studie hinsichtlich wesentlicher demographischer Faktoren charakterisiert.



**Abbildung 6.8.:** Visuelle Zusammenfassung der wichtigsten Vorverarbeitungsschritte

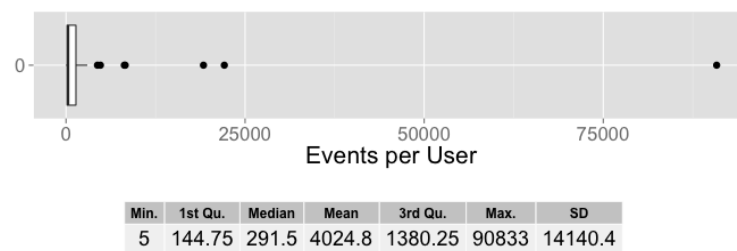
### 6.4.2. Charakterisierung der Studienteilnehmer

Im Studienzeitraum konnten nach Bereinigung und Mapping von Nutzer-IDs 44 Teilnehmer identifiziert werden. Aufgrund der Tatsache, dass Rekrutierungsstrategien von unterschiedlichem Erfolg gekrönt waren, zu unterschiedlichen Zeitpunkten initialisiert wurden und nicht alle Teilnehmer gleich häufig Twitter über die Website nutzen, waren nicht alle Teilnehmer über die gesamte Studiendauer aktiv.



**Abbildung 6.9.:** Boxplot (a) und Tabelle (b) zur Darstellung der Verteilung der Studienteilnahmedauer in Tagen

Abbildung 6.9 zeigt, dass die Teilnehmer im Durchschnitt 55,44 Tage an der Studie teilgenommen



**Abbildung 6.10.:** Boxplot (a) und Tabelle (b) der Anzahl Ereignisse pro Studienteilnehmer

haben, wobei ein Viertel aller Teilnehmer mindestens 3 Monate ( $3.Quantil = 102$ ) oder länger Interaktionsdaten geliefert hat. Es lässt sich erkennen, dass die Zeitspanne an aktiver Teilnahme von ProbandIn zu ProbandIn sehr unterschiedlich war. Neben den bereits genannten Ursachen lässt sich dies auch mit dem Umstand begründen, dass für viele Nutzer das Smartphone oder Tablet und somit die mobile Nutzung die Hauptnutzungssituation darstellen. In Abschnitt 5.5 — im Zuge der Darstellung der demographischen Eigenschaften der beiden Stichproben der Twitter-Umfrage — konnte jedoch festgestellt werden, dass Nutzer häufig sowohl die Website als auch eine mobile Twitter-App nutzen. Innerhalb des Studienzeitraums erzeugten 44 Teilnehmer 177091 Interaktionsereignisse, wobei hier einzelne Mausinteraktionen ausgenommen sind. Einzelne Scrolling-Ereignisse werden bei einer Pause der Mausbewegung zu einem `scroll summary`-Ereignis zusammengefasst und sind in der Gesamtzahl enthalten. Die Verteilung der Events auf die einzelnen Nutzer der Erweiterung fällt noch extremer aus. Abbildung 6.10 zeigt die Verteilung der Events pro Studienteilnehmer. Der Median liegt bei 291,5 Ereignissen pro Teilnehmer, was offenbart, dass die Hälfte aller Teilnehmer nur sehr wenige Ereignisse generierten und der Großteil der 177091 Ereignisse von einer geringen Anzahl an Probanden erzeugt wurde. Ein Proband ist mit 90833 (Max) Events für über die Hälfte (51,30%) aller Ereignisse verantwortlich.

Von 44 Studienteilnehmern konnten 27 zu einem bekannten Twitter-Account zugeordnet werden.<sup>24</sup> Die Account-Informationen sind in Tabelle 6.2 aufgeführt. Die Teilnehmer der Studie bilden eine sehr heterogene Gruppe, da sowohl Werte, die als Platzhalter für aktive Teilnahme interpretiert werden können, wie gesendete Tweets ( $Min = 0$ ,  $Max = 73780$ ) und die Anzahl an favorisierten Tweets ( $Min = 0$ ,  $Max = 44520$ ), eine sehr breites Spektrum an unterschiedlicher Nutzungsintensität aufzeigen als auch in Bezug auf die Größe des sozialen Netzwerkes (Anzahl der Followees und Anzahl der Follower) starke Unterschiede auszumachen sind. Die Werte für die Jahre, seit denen Nutzer bei der Plattform registriert sind, sind nur bedingt aussagekräftig, da hier keine Pflichtangabe besteht und daher von zwölf Nutzern der Wert nicht bekannt ist. Die restlichen Teilnehmer sind mit einem Median von fünf Jahren durchaus als Langzeitnutzer zu betrachten.

Um die Anonymität der Studienteilnehmer, welche im Kontext einer Logstudie stets ein gewisses ethisches Problem darstellt, so gut wie möglich zu gewährleisten, wurden von den Studienteilnehmer keine weiteren demografischen Informationen eingeholt. Durch Begutachtung ihrer Profile konnte jedoch festgestellt werden, dass die Probanden in unterschiedlichen deutsch- und englischsprachigen Ländern beheimatet sind (hauptsächlich kamen die Teilnehmer aus Deutschland, England und den USA) sowie in unterschiedlichen Beschäftigungsverhältnissen stehen.

Vor Abschluss der Studie wurden die Teilnehmer in einem finalen Fragebogen dazu befragt, wie

<sup>24</sup>Die Identifikation von Nutzern wurde über den Besuch des eigenen Profils während des Studienzeitraums ermittelt. Falls kein Profilbesuch getätigt wurde, wurde versucht, intellektuell ein Mapping (ErweiterungsID - Twitter-Account) herzustellen, wie bereits in Abschnitt 6.4.1 beschrieben.

(n=27)	Min	Max	Mittelwert	Median
Tweets gesendet(n)	0	73 780	5637	59
Anzahl Folge ich(n)	4	1041	203	57
Anzahl Follower(n)	0	1139	202,50	39
Anzahl Favoriten(n)	0	44520	1796	18,50
Registriert seit (Jahre) (NA=12)	2	6	4,60	5

**Tabelle 6.2.:** Twitter-Account-Informationen der Studienteilnehmer

häufig sie während des Studienzeitraums Twitter über die Twitter-Website genutzt haben und wie häufig andere Geräte oder Applikationen hierfür genutzt wurden. Nur 14 Versuchsteilnehmer haben diesen finalen Fragebogen beantwortet, was abermals zeigt, dass nicht alle Teilnehmer Twitter kontinuierlich über die Website nutzen bzw. über die volle Studienzeit hinweg aktiv an der Studie teilgenommen haben. Dabei geben vier Probanden (28,57%) an, mehrfach am Tag oder täglich Twitter über den Website-Client genutzt zu haben. Ebenso viele Probanden (28,57%) geben an, mehrfach die Woche Twitter über die Website genutzt zu haben. Zwei Probanden (14,29%) nutzen die Twitter-Website wöchentlich, wohingegen 21,43% der Teilnehmer die Website mehrmals im Monat nutzten und 14,29% der Teilnehmer dies nur selten tun. Als Antwort auf die Frage, welche andere Geräte und Applikationen zur Twitter-Nutzung verwendet werden, geben 71% an, die je nach Betriebssystem (Android/iOS) für mobile Geräte entwickelte Twitter-App zu verwenden. Zwei der Probanden geben an, nur die Website über Chrome zu nutzen. Ein Proband verwendet zusätzlich Tweetdeck. Ein weiterer Proband den mobilen Chrome Browser. Interessanterweise geben die Probanden nicht an, diese Zugangsweisen weitaus häufiger zu nutzen als die Twitter-Website. Nur 2 Probanden berichten, die Twitter-App am Smartphone täglich zu nutzen und damit deutlich häufiger als die Website. Diese Erkenntnisse zum Nutzungsverhältnis decken sich mit den Erkenntnissen aus Abschnitt 5.5.

Ein statistisch repräsentatives Sample an Twitter-Nutzern für eine Logstudie dieser Art zu rekrutieren, ist ein nahezu unmögliches Unterfangen. Die hierfür benötigten Daten/Ressourcen würden wahrscheinlich nur von Twitter selbst aufgebracht werden können. Das Ziel dieser Studie ist es, Re-finding-Verhalten so objektiv wie möglich mit einer heterogenen, divers charakterisierten Stichprobe zu analysieren. Betrachtet man die Beschreibung der Studienteilnehmer in den letzten Abschnitten, wird deutlich, dass das Spektrum an unterschiedlichen Nutzungs- und Verhaltensweisen sehr breit ist. Dennoch ist eine gewisse Befangenheit in Bezug auf wenige, sehr aktive Nutzer nicht von der Hand zu weisen. Trotzdem konnte ein diverses Sample rekrutiert werden, das sich gut zur Analyse der Fragestellungen eignet, wie in Abschnitt 6.4.4 gezeigt wird.

### 6.4.3. Twitter-Verhalten im Allgemeinen

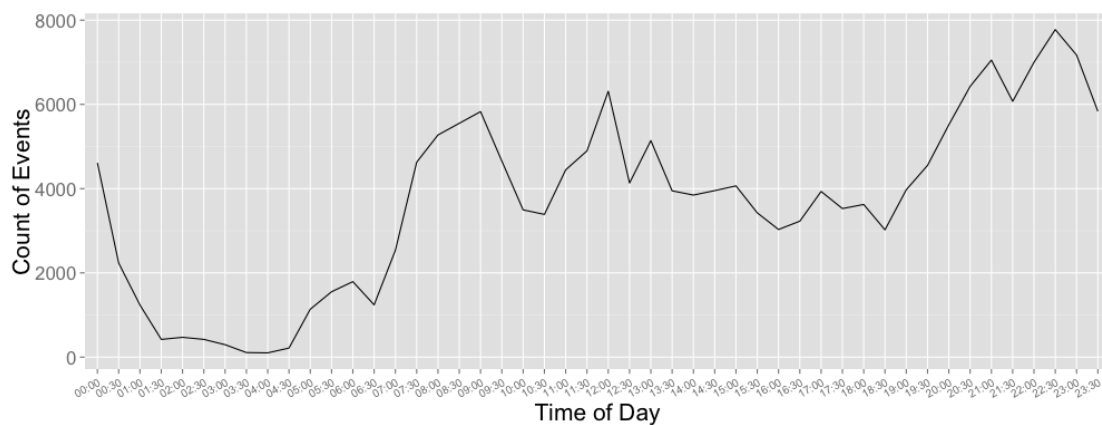
Zunächst wird die Leitfrage nach dem allgemeinen Twitter-Verhalten behandelt. Es existieren gewissermaßen keine Studien, die sich mit genereller Twitter-Nutzung beschäftigen, also aufzeigen, welche Nutzungsszenarien es gibt, wodurch diese charakterisiert sind, welche Gründe oder Motive eine Twitter-Nutzung motivieren, wie lange am Stück Nutzer aktiv sind, oder welche Informationsbedürfnisse sie zu befriedigen versuchen. Lediglich Russell-Rose, Lamantia und Makri erläutern anhand des Twitter-User-Interfaces, welche vermeintlichen Nutzungsmöglichkeiten — charakterisiert durch sogenannte Sucharten und deren Verkettung — Nutzern zur Verfügung stehen (Russell-Rose et al., 2014). Dabei werden aber keine Aussagen darüber getroffen, wie häufig unterschiedliche Verhaltensweisen wie die sogenannte *strategic oversight chain* (als Kette aus den Suchstrategien *mo-*



nitor — analyze — evaluate) an den Tag gelegt werden, sondern es werden lediglich UI-Features präsentiert, die die einzelnen Strategien unterstützen (Russell-Rose et al., 2014, S.15). Somit wird vor der Darstellung des Re-finding-Verhaltens, welches ein sehr spezifisches Nutzungsszenario darstellt, das Twitter-Verhalten der Nutzer im Allgemeinen, d.h. ohne spezifischen Fokus, beleuchtet.

Zu Beginn stellt sich die Frage, zu welcher Tageszeit die Teilnehmer der Studie Twitter hauptsächlich genutzt haben. Abbildung 6.11 zeigt die Twitter-Nutzung der Studienteilnehmer im Tagesverlauf. Es zeigt sich eine klassische, an den Tagesrhythmus angepasste Kurve. In der Nacht, zwischen 00:00 und 06:00, fällt die Twitter-Nutzung auf ein Minimum ab, bevor sie morgens wieder zu steigen beginnt. Über den Tagesverlauf hinweg bleibt die Nutzung etwa konstant bzw. stets ansteigend und ist von drei Höhepunkten um 9:00 Uhr, zur Mittagszeit um 12:00 Uhr und einem deutlichen Höhepunkt um 22:30 Uhr, gekennzeichnet. Ähnliche Nutzungsweisen finden sich bei Social-Media-Loganalysen von Plattformen wie *Renren* (G. Wang et al., 2013) oder *Facebook* (Schneider et al., 2009).

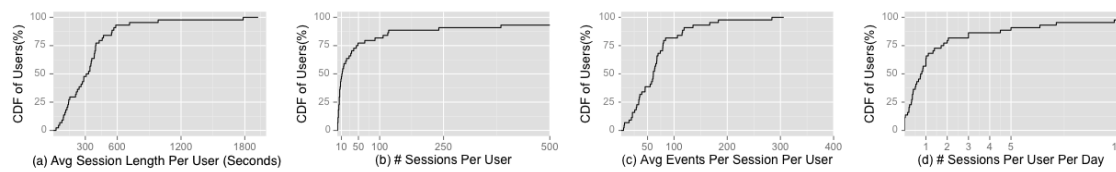
#### 6.4.3.1. Twitter-Verhalten in Bezug auf Sessions



**Abbildung 6.11.:** Twitter-Nutzung der Studienteilnehmer im Tagesverlauf

Die Wahl einer fünfminütigen Inaktivität als Sessiongrenze teilt die Logdaten in 4573 Sessions ein. Die Anzahl an Ereignissen pro Session beträgt im Durchschnitt 38,7 wobei auch hier starke Unterschiede auszumachen sind, da die Anzahl zwischen minimal 2 und maximal 2155 Events pro Session variiert ( $Median = 11$ ,  $IQR = 31$ ). Eine noch extremere Verteilung zeigt sich bei der Dauer einer Session, welche zwischen 1 Sekunde und 3 Stunden (11080 Sekunden) dauern kann. Im Durchschnitt dauerte eine Twitter-Session der Studienteilnehmer etwas über 4 Minuten ( $Median = 83$  Sek.,  $IQR = 270$  Sek.). Die Anzahl der Events pro Session und Sessiondauer korrelieren äußerst stark (Pearson  $\rho = 0,90$ ), wodurch deutlich wird, dass Twitter-Nutzer sehr aktiv sind und wenig reine Rezeptionsphasen vorhanden sind, in denen keine Interaktion erfolgt.

Die Abbildungen 6.12 a) - d) vermitteln einen Eindruck über das generelle Sessionverhalten der Studienteilnehmer in Form von empirischen Verteilungsfunktionen. Dabei ist anzumerken, dass je steiler die Kurve verläuft, desto extremer die Verteilung der Daten ist und sich ein sogenannter *long tail* zu erkennen gibt. Am stärksten ist dies bei Abbildung b) zu sehen, wobei deutlich wird,



**Abbildung 6.12.:** Charakterisierung der Studienteilnehmer in Bezug auf Sessioncharakteristika

dass die Anzahl an Sessions pro Nutzer extrem unterschiedlich ist und eine Vielzahl an Probanden nur sehr wenige Sessions aufzuweisen hat. 75% aller Teilnehmer haben Twitter in 46 oder weniger Sessions genutzt. Drei Probanden besitzen jedoch an die 1000 aktive Sessions. Abbildung 6.12 d) zeigt, dass etwa die Hälfte (*Median* = 0,8) aller Nutzer Twitter einmal pro Tag benutzt hat. Auch hier ist die Spanne der Nutzungsweise sehr divers. Ein Viertel (*1.Quantil* = 0,3) aller Teilnehmer durchstöbert ihre *Timeline* nur jeden 3. Tag, jedoch ruft eine Person ihren Account auch etwa 12 (*Max* = 12,5) mal pro Tag auf.

Die Anzahl der Sessions muss auch in Relation zur durchschnittlichen Anzahl an Events pro Session betrachtet werden, da einerseits auch Personen mit wenigen Sessions sehr intensiv mit Twitter interagieren können (hohe durchschnittliche Anzahl an Events pro Session), andererseits eine Vielzahl an Einzelsessions nicht mit intensiver Twitter-Nutzung gleichzusetzen ist (wenige Events pro Session). Im Median besitzen Studienteilnehmer 61 Events pro Session. Das Maximum liegt hier bei 284 Events. Abbildung 6.13 zeigt die Anzahl an Sessions pro Nutzer und die durchschnittliche Anzahl an Events pro Session pro Nutzer im Verhältnis. Es wird abermals deutlich, dass viele Nutzer wenige Sessions besitzen, die durchschnittliche Anzahl an Events dabei aber sehr stark variiert. Jedoch wird auch deutlich, dass einige wenige Nutzer sehr häufig mit Twitter interagieren, dabei aber nur sehr kurz bzw. sehr wenige Events generieren, was abermals sehr unterschiedliches Nutzungsverhalten aufzeigt. Im Folgenden wird Twitter-Verhalten auf der Basis der in Abschnitt 6.4.1.1 eingeführten Ereigniskategorien beschrieben.

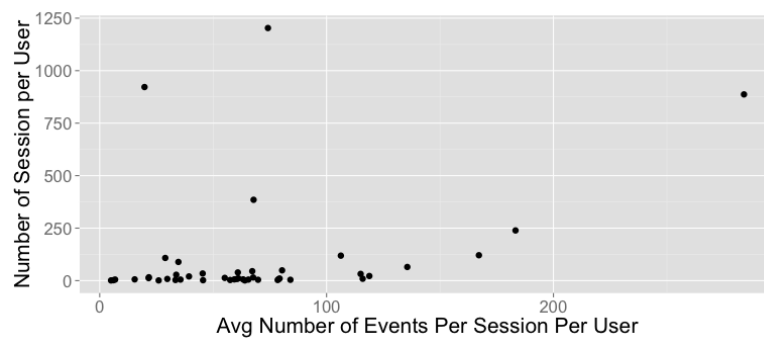
#### 6.4.3.2. Twitter-Verhalten in Bezug auf Eventkategorien

Um zu überprüfen, ob das Auftreten bestimmter Events bzw. Eventkategorien mit der Dauer der Nutzung in Zusammenhang steht, wurden Sessions in Anlehnung an die Vorgehensweise von Benevenuto und Kollegen basierend auf ihrer Dauer in vier Gruppen eingeteilt und jeweils der Anteil jeder Eventkategorie an den Gesamtereignissen pro Zeitgruppe berechnet (Benevenuto et al., 2009). Abbildung 6.14 zeigt die Verteilung von Ereignissen auf die jeweilige Ereigniskategorie im Verhältnis zur Gesamtanzahl der Ereignissen bei vier unterschiedlichen Sessionlängen von sehr kurz (bis 1 min) bis sehr lang (länger als 20 min). Da manche Eventkategorien einen deutlich größeren Anteil an den Gesamtereignissen ausmachen — die meisten Ereignisse finden im Kontext *Timeline* statt — wurde die Grafik auf dominante (links) und eher untergeordnete (rechts) Eventkategorien bzw. Ereignisse aufgeteilt.

##### Dominante Aktivitäten

Wie bereits erwähnt, finden die meisten Ereignisse (72,4%) auf der eigenen *Timeline*<sup>25</sup> statt. Vergleicht man den Anteil an Ereignissen in der Kategorie *Timeline* über die vier Zeitgruppen hinweg, lässt sich erkennen, dass der Anteil an Events in kurzen bis mittellangen Sessions (> 1 min < 10 min) um etwa 15% abnimmt, dann aber wieder bis zu einem Ursprungswert von etwa 80% steigt.

<sup>25</sup>Die eigene *Timeline* entspricht der URL bzw. dem Kontext: <https://twitter.com>



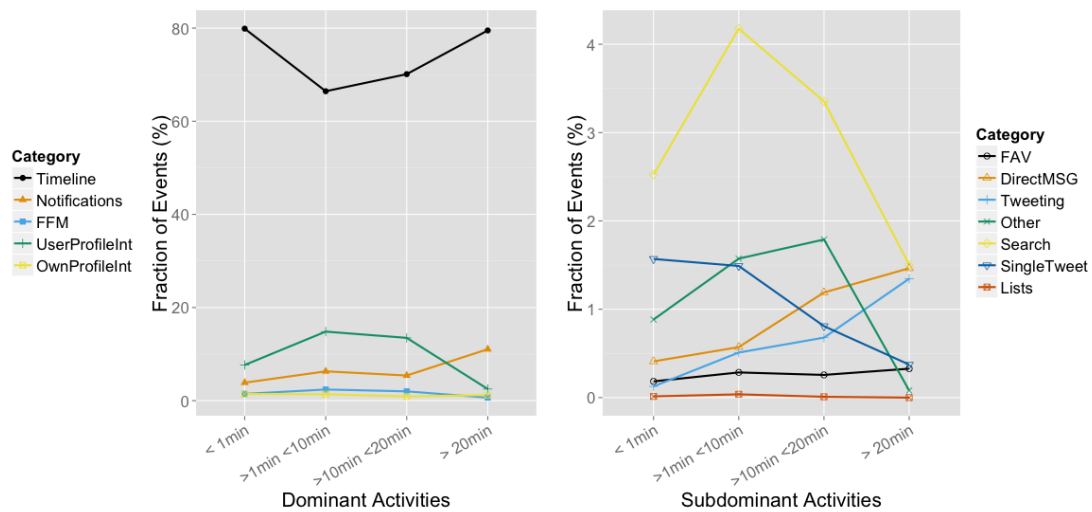
**Abbildung 6.13.:** Anzahl an Sessions pro Nutzer im Verhältnis zu dessen durchschnittlicher Anzahl an Events pro Session

Der zweithäufigste Kontext, in welchem Ereignisse stattfinden, sind Nutzerprofile von anderen Twitter-Nutzern (Kontext: `USERPROFILEINT`). 10,1% aller Ereignisse entfallen auf diese Eventkategorie. In kurzen bis mittellangen Sessions ( $> 1 \text{ min} < 10 \text{ min}$ ) lässt sich eine Verdopplung der Nutzerprofilereignisse von 7,7% auf 15% im Verhältnis zu sehr kurzen Sessions ( $< 1 \text{ min}$ ) beobachten. Die Anzahl an Ereignissen im Kontext `NOTIFICATIONS` nimmt kontinuierlich über die Sessionlänge hinweg zu. Es wird deutlich, dass sehr kurze Sessions stark durch *Timeline*-Interaktionen geprägt sind, was vermutlich eine Konsequenz aus dem Verhalten ist, welches am besten als *Überprüfen von neuesten Tweets* beschrieben werden kann. Weniger offensichtlich hingegen ist die starke Zunahme an Nutzerprofilevents, die - zumindest in mittellangen Sessions - einen sehr großen Anteil an den Gesamtevents einnehmen. Besuche des eigenen Profils (`OWNPROFILEINT`) oder Ereignisse, die mit der Kategorie `FFMRELATIONS` in Zusammenhang stehen, machen nur einen sehr geringen Anteil der Interaktion aus und treten im Verhältnis also eher selten auf.

#### Untergeordnete Aktivitäten

Betrachtet man die eher seltener auftretenden Ereignisse in Abbildung 6.14 (rechts) wird deutlich, dass das Senden von eigenen Tweets oder Ereignissen, die mit Direktnachrichten in Beziehung stehen (Senden/Lesen), wahrscheinlicher werden, je länger eine Session dauert. Liegt der Anteil an Events aus der Kategorie `TWEETING` bei sehr kurzen Sessions noch bei 0,1% steigt dieser auf bis zu 1,4% bei sehr langen Sitzungen ( $> 20 \text{ min}$ ). Das schnelle, spontane Absenden eines Tweets, welches nicht durch weitere Interaktionen begleitet ist, scheint im Nutzungskontext Website kaum zu existieren. Eine vergleichbare Entwicklung nimmt das Lesen bzw. Senden von Direktnachrichten, welche von 0,4% ( $< 1 \text{ min}$ ) auf 1,5% ( $> 20 \text{ min}$ ) steigt. Nur 3,1% aller Ereignisse gehören zur Eventkategorie `SEARCH`. Interessant ist dabei die Entwicklung, welche die Kategorie über die vier Sessiongruppen nimmt. Spielen Ereignisse der Kategorie `SEARCH` in sehr kurzen Sitzungen eine eher untergeordnete Rolle, sind 4,2% aller Events in kurzen bis mittellangen Sessions ( $> 1 \text{ min} < 10 \text{ min}$ ) Suchereignisse. In längeren Sitzungen ( $> 10 \text{ min} < 20 \text{ min}$  und  $> 20 \text{ min}$ ) nimmt der Anteil an Suchereignissen aber deutlich ab und fällt auf 1,5% aller Ereignisse. Twitter-Suchersitzungen scheinen eher von kurzer Dauer zu sein. Gründliche, langandauernde Suchsessions gehören somit eher der Seltenheit an. Schließlich kann man erkennen, dass die Probanden ihre Favoritenliste nur selten besuchen. Nur 0,3% aller Ereignisse sind Ereignisse aus der Kategorie `FAVOURITES`. Lediglich 8 Nutzer (18,2%) betrachten ihre Favoriten während des Studienzeitraums. Die Listenfunktion wurde von den Studienteilnehmern in der Logstudie kaum genutzt.

Auf ähnliche Weise kann beobachtet werden, ob das Auftreten bestimmter Ereigniskategorien an gewisse Tageszeiten gebunden ist. Abbildung 6.15 zeigt den Anteil von Eventkategorien an den Gesamtereignissen über den Tagesverlauf verteilt. Wie bei der Darstellung in Abbildung 6.14 wird



**Abbildung 6.14.:** Anteil der Ereigniskategorien an den Gesamtereignissen kategorisiert in Gruppen nach Sessiondauer

zwischen dominanten und eher seltener auftretenden Aktivitäten unterschieden.

#### Dominante Aktivitäten

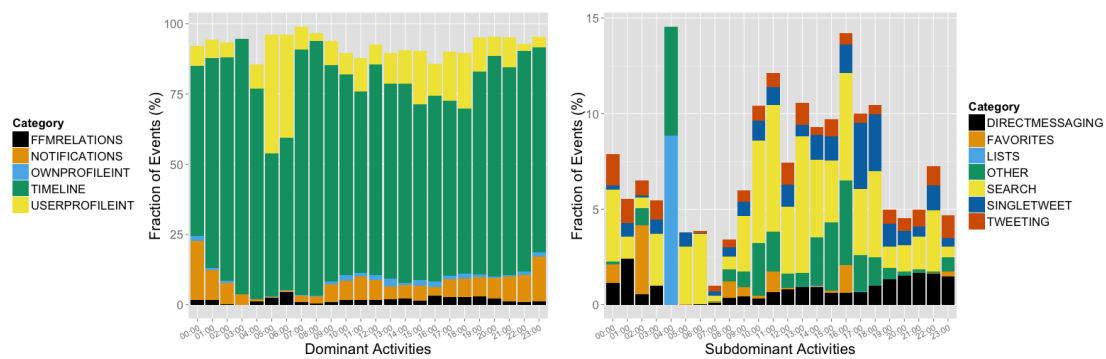
In Bezug auf die Kategorie *Timeline* lassen sich wenig Auffälligkeiten ausmachen. Sie stellt stets den dominanten Anteil dar. Lediglich zwischen 5:00 und 6:00 Uhr morgens geht der Anteil der Events aus der Kategorie *Timeline* auf 51% bzw. 54,4% der Gesamtereignisse zurück und Ereignisse auf Profilen von anderen Nutzern (*USERPROFILEINT*) steigen auf ihre Höchstwerte von 42,3%, respektive 36,6% an. Diese Werte müssen im Kontext der Abbildung 6.11 betrachtet werden, die zeigt, dass zu dieser Uhrzeit ohnehin sehr wenig Events auftreten. Eine weitere Auffälligkeit zeigt sich bei der Event-Kategorie *NOTIFICATIONS*, welche über die Mittagszeit und am Abend einen stärkeren Anstieg erfährt und mit 21,00% ihren Höhepunkt zwischen 23:00 und 00:00 findet. Bei den dominanten Aktivitäten lassen sich sonst keine weiteren Auffälligkeiten oder Besonderheiten feststellen.

#### Untergeordnete Aktivitäten

In den seltener auftretenden Ereigniskategorien zeigt sich vor allem in Bezug auf *SEARCH*, die Betrachtung einzelner Tweets (*SINGLETWEET*) und das Absenden von Tweets (*Tweeting*) ein einheitliches Muster, so dass keine Aussage darüber getroffen werden kann, dass eine dieser Kategorien zu einem bestimmten Tageszeitpunkt häufiger auftreten würde. Folgende drei Besonderheiten sind jedoch in Bezug auf den Tagesverlauf zu bemerken:

1. Um 4 Uhr nachts findet sich ein hoher Anteil an untergeordneten Events im Vergleich zu anderen Tageszeitpunkten. Dabei ist besonders auch der hohe Anteil an Ereignissen der Kategorie *List* (8,7%) auffällig.
2. Um 7 Uhr morgens ist der Anteil an dominanten Aktivitäten äußerst hoch, wodurch nur etwa 1% auf subdominante Aktivitäten entfällt.
3. Je später am Abend die Nutzer mit Twitter interagieren, desto häufiger werden auch Direktnachrichten gelesen oder gesendet. Der Anteil an Ereignissen der Kategorie *DIRECTMESSAGE* ist morgens noch gering und steigt über den Tagesverlauf hinweg kontinuierlich an.

Zusammenfassend lässt sich feststellen, dass bis auf wenige Ereigniskategorien wie das Betrachten der *Notifications*-Seite und die Nutzung der Direktnachrichtenfunktion — welche beide über den



**Abbildung 6.15.:** Anteil der Ereigniskategorien an den Gesamtereignissen zu festen Tageszeitpunkten

Tagesverlauf zunehmen und am Abend am häufigsten genutzt werden — keine Zusammenhänge zwischen Tageszeit und Eventkategorie auszumachen sind.

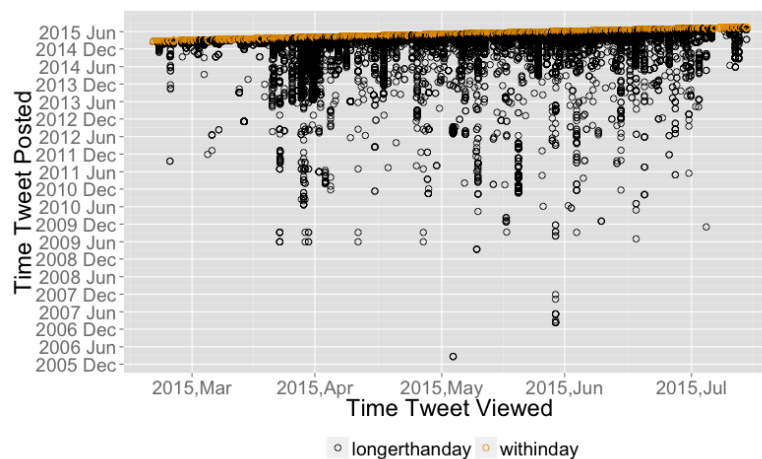
### 6.4.3.3. Twitter-Verhalten in Bezug auf Interaktionen mit Tweets

Im Studienzeitraum wurden 143 721 einzelne Tweets (*unique Tweets*) aufgezeichnet, wobei 273 550 Referenzen auf diese Tweets (*collected Tweets*) registriert wurden. 146 735 dieser Referenzen sind eindeutig pro Studienteilnehmer (*unique Tweets per User*), was bedeutet, dass ein Tweet zweimal zu dieser Statistik gewertet wird, wenn er in den *Timelines* von zwei unterschiedlichen Nutzern erscheint. Lediglich bei 3,7% dieser Tweets konnten twitterspezifische Interaktionen festgestellt werden. 3% der Tweets wurden favorisiert (FAV), 0,3% wurden an die eigenen Follower weitergeleitet (RT) und 0,4% wurden sowohl favorisiert als auch weitergeleitet (FAV+RT). Starke Unterschiede im Nutzungsverhalten der Studienteilnehmer zeigen sich dadurch, dass diese Interaktionen nur von 23 (52,3%) Personen getätigt werden. Zudem kann Interaktionsverhalten sehr einseitig charakterisiert sein. 83% (4186 Tweets) aller Favorisierungen wurden von einem Probanden getätigt, der dieses Feature exzessiv nutzt.

Eine in Abschnitt 6.1 formulierte Leitfrage stellt die Frage nach dem nutzerbezogenen, individuellen Lebenszyklus von Tweets. Das Ziel der folgenden Abschnitte ist die Untersuchung, inwiefern Tweets einen verlängerten Lebenszyklus besitzen, respektive auch nach längerer Zeit — seit ihrer initialen Veröffentlichung — für einen Nutzer von Wert/Bedeutung sein können. Einen wesentlichen Faktor spielt hierbei das Alter eines Tweets. Abbildung 6.16 zeigt Veröffentlichungszeitpunkt und Betrachtungszeitpunkt der Tweets durch die Studienteilnehmer im Verhältnis. Die Kurve zeigt einen leichten Anstieg, was der Tatsache geschuldet ist, dass bei fortschreitender Studiendauer Tweets immer jünger werden bzw. der Betrachtungszeitpunkt immer näher an die aktuelle Zeit annähert und sich die Linie dem rechten oberen Eck annähert.

Das Durchschnittsalter eines Tweets beträgt zum Betrachtungszeitpunkt 28,1 Std.

(*Median* = 1,4 Std., *Max* = 80667 Std. (9 Jahre), *IQR* = 11 Std.). Zudem ist farblich hervorgehoben (orange Punkte), ob Tweets innerhalb von 24 Stunden seit ihrer Veröffentlichung betrachtet werden, oder erst zu einem späteren Zeitpunkt. Obwohl die meisten Tweets sehr aktuell sind — 80,9% aller Tweets werden innerhalb von 24 Std. seit ihrem Posting gesehen — sind einige bereits mehrere Stunden oder sogar Tage alt, wenn sie vom Nutzer betrachtet werden. Diese Tatsache wird deutlich in Abbildung 6.16, in der zu erkennen ist, dass manche Tweets (19,1%) erst nach 24 Stunden ihrer Veröffentlichung betrachtet werden, respektive bereits einige Jahre alt sind, wenn



**Abbildung 6.16.:** Veröffentlichungs- und Betrachtungszeitpunkte der gesammelten Tweets

sie am Bildschirm des Nutzers erscheinen. Studienteilnehmer interagieren mit Tweets, die bis ins Jahr 2006 zurückreichen. Sogar der Ur-Tweet, welcher von Jack Dorsey am 21. März 2006 gesendet wurde (vgl. Abbildung 6.17), ist in den Logs vermerkt. All dies lässt den Schluss zu, dass die Charakterisierung Twitters als Echtzeit-Informationsnetzwerk nur bedingt zutrifft und Twitter auch einen informationsbewahrenden Charakter besitzt. Das hohe Alter, das manche Tweets zum Betrachtungszeitpunkt vorweisen, ist ein erstes Indiz dafür, dass Tweets einen erweiterten Lebenszyklus besitzen und nach Jahren ihrer Veröffentlichung noch von Relevanz sein können (siehe Tweet Jack Dorsey).<sup>26</sup>

Anhand der Darstellung der Sende- und Betrachtungszeitpunkte von Tweets, lässt sich auch abermals das unterschiedliche Nutzungsverhalten der Studienteilnehmer deutlich machen. Abbildung 6.18 vergleicht zwei Nutzer hinsichtlich des Alters der Tweets, die sie bei ihren Sitzungen betrachtet haben. Zum einen wird deutlich, dass der Studienteilnehmer mit der ID 14 Twitter weitaus kontinuierlicher genutzt hat, was an der fast durchgehenden Linie zu erkennen ist, welche die orangenen Punkte bilden. Bestätigt wird dies durch den relativ hohen Prozentsatz an Tweets (71%), welche innerhalb von 24-Stunden betrachtet werden. Zum anderen ist zu erkennen, dass die Skala der Y-Achse nicht weiter zurückreicht als in den Juni 2014, was deutlich macht, dass kaum ältere Tweets betrachtet wurden. In Kontrast hierzu wird bei der Darstellung desselben Plots von Studienteilnehmer mit ID 57 deutlich, dass dieser Twitter weitaus seltener über das Web-Interface benutzt und mehrfach auch ältere Tweets betrachtet werden (*Durchschnittsalter* = 465,3 Std.). Dennoch wird auch hier die überwiegende Mehrheit an Tweets (75,2%) innerhalb eines Tages betrachtet. Das Alter eines Tweets wird in Abschnitt 6.4.4.2 nochmals diskutiert, wenn unterschiedliche Varianten von *Proxies* d.h. Ersatz- oder Platzhaltermaße zur Erkennung von Re-finding-Verhalten diskutiert werden.

Neben dem Tweetalter ist auch die Position eines Tweets in der *Timeline* von Bedeutung, da ältere Tweets tendenziell weiter unten in der *Timeline* zu finden sind und die Position darüber Auskunft gibt, wie viel Scrolling-Aufwand der Nutzer auf sich genommen hat, um entsprechende Tweets auf dem Bildschirm erscheinen zu lassen.<sup>27</sup> Der Median der Tweetposition liegt bei 7 (*Min* = 0,

<sup>26</sup>Es ist anzumerken, dass durch die invertierte chronologische Anordnung der Tweets in der *Timeline* stets die aktuellsten Tweets gezeigt werden und nur dann ein Bias in Bezug auf Nutzer besteht, die nur sehr selten Twitter nutzen, wenn diese zeitgleich nur sehr wenigen Nutzern folgen, respektive diese Nutzer auch eine sehr geringe Tweet-Frequenz besitzen.

<sup>27</sup>Dabei ist die Position des Tweets noch stärker von der Anzahl an Freunden und deren Tweetfrequenz abhängig, denn je



**Abbildung 6.17.:** Der allererste Tweet gesendet von Jack Dorsey am 21. März 2006 )

(Screenshot von: <https://twitter.com/jack/status/20>)

$Mean = 38,04$ ,  $IQR = 41$ ), das Maximum bei 627. Durch das hohe Maximum wird erkenntlich, dass Nutzer in bestimmten Situation sehr viel Aufwand betreiben, um an bestimmte Tweets zu gelangen. Wodurch dieses Verhalten motiviert ist, bleibt zu diesem Zeitpunkt noch offen. Jedoch kann die hohe Position und der damit verbundene Aufwand ein weiteres Indiz für Re-finding sein.

#### 6.4.3.4. Zusammenfassung und Erkenntnisse für die Re-finding-Analyse

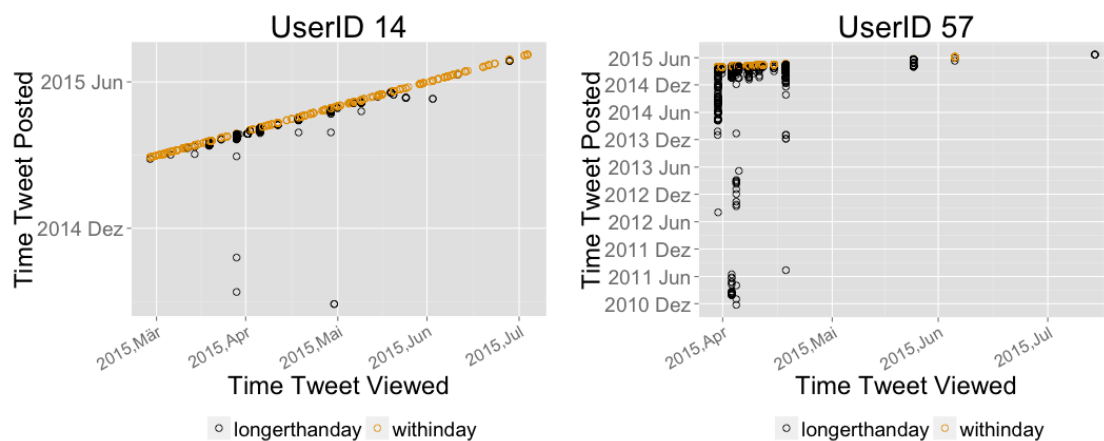
Aufgrund der Tatsache, dass es keine Vorarbeiten zur generellen Twitter-Nutzung gibt und der Einzigartigkeit des Datensatzes sind die Erkenntnisse zur Nutzung in Bezug auf Sessions, den Unterseiten bzw. Ereigniskategorien sowie dem Umgang mit Tweets per se von hoher Relevanz. Letztlich stellt sich aber die Frage, welche Konsequenz diese Erkenntnisse für die Hauptfragestellung des Wiederfindensverhaltens besitzen. Interessanterweise lassen sich bereits bei der Betrachtung des generellen Twitter-Nutzungsverhaltens einige Indizien bzw. Hinweise auf Re-finding-Verhalten ausmachen.

- Das durchschnittliche Alter gesichteter Tweets ist mit 28,1 Std. unerwartet hoch. Auch die Tatsache, dass Tweets, die bis zu 9 Jahre alt sind in den Logs vertreten sind, überrascht. Beide Umstände zeigen, dass Tweets einen verlängerten Lebenszyklus haben könnten, und nach einer gewissen Zeit erneut von Interesse für den Nutzer sind, weswegen sie wiedergefunden werden.
- Die Position von Tweets, mit denen interagiert wird, liegt im Median bei Rang 7. Analog zum Tweetalter zeigen sich auch bei der Position extreme Werte ( $Max = 627$ ), was darauf hindeutet, dass Nutzer sehr viel Aufwand betreiben, um zu älteren Tweets zu gelangen: Erneut ein möglicher Hinweis für Re-finding-Verhalten.
- Eventkategorien/Unterseiten, die laut Abschnitt 3.2.2.2 stark mit Speicherung/Verwaltung und Wiederfinden von Tweets in Zusammenhang stehen (FAVOURITES und OWNPROFILEINT) werden überraschenderweise selten besucht. Wenn also Wiederfindensverhalten stattfindet, dann ist dieses Verhalten anders charakterisiert als in der Literatur beschrieben (boyd et al., 2010).

Im Folgenden wird Wiederfindensverhalten in den Logdaten identifiziert und analysiert sowie die eingangs gestellten Hauptleitfragen bearbeitet.

---

mehr Nutzern man folgt und je mehr Tweets diese produzieren, desto schneller rücken auch sehr aktuelle Tweets in der *Timeline* weiter nach unten.



**Abbildung 6.18.:** Veröffentlichungs- und Betrachtungszeitpunkte der Tweets zweier Nutzer im Vergleich. Die Grafiken zeigen unterschiedliches Nutzungsverhalten.

#### 6.4.4. Analyse des Re-finding-Verhaltens

Um Wiederfindensverhalten analysieren zu können, müssen zunächst Re-finding-Aktionen in den Logs identifiziert und extrahiert werden. Für die Identifikation von Re-finding-Verhalten in Logdaten werden in der Literatur unterschiedliche Verfahren beschrieben. Capra stellt in einer Laborstudie mit 18 Suchaufgaben fest, dass es schwierig ist, zwischen normalem Suchmaschinengebrauch und Wiederfindensverhalten zu unterscheiden (Capra & Pérez-Quinones, 2003). Elsweiler und Kollegen nutzen einen manuell annotierten Goldstandard, um ein logistisches Regressionsmodell zu trainieren, das anhand von heuristischen Features E-Mail-Sessions als Re-finding-Sitzungen klassifiziert (Elsweiler, Harvey & Hacker, 2011). Sadeghi et al. arbeiten ebenfalls mit maschinellen Lernverfahren und nutzen einen Klassifikator, um Session-Paare zu bilden, die die gleichen Ziele besitzen (Sadeghi et al., 2015). Ähnliche Verfahren werden auch verwendet, um Suchaufgaben zu erkennen, die sich über mehrere Sessions hinweg verteilen (Kotov et al., 2011). Sehr häufig werden Interaktionssequenzen — erneute Aktionen über Sessiongrenzen hinweg — als Ersatzmaße für Wiederfindensverhalten gewertet. In einer ersten Arbeit von Teevan und Kollegen, die auf Logdatenanalyse zur Charakterisierung von Re-finding-Verhalten bei der Websuche setzen, wird ein erneuter Klick auf das gleiche Suchmaschinenergebnis als Re-finding-Absicht gewertet (Teevan, Adar et al., 2007). Auch bei Tyler und Teevan sowie bei Harvey und Elsweiler wird eine sessionübergreifende Klicksequenz als Re-finding-Absicht interpretiert (Tyler & Teevan, 2010; Harvey & Elsweiler, 2012). Tyler und Teevan betrachten jedoch auch die erneute Interaktion mit einem Suchmaschinenergebnis innerhalb einer Session als Re-finding (Tyler & Teevan, 2010). Dies ist jedoch eher als *Re-visit* zu interpretieren. Ein *Re-visit* im vorliegenden Kontext würde bedeuten, dass ein Tweet mehrfach innerhalb einer Sitzung geklickt wird. Dieses Verhalten wird jedoch in der Arbeit nicht näher betrachtet. Im Folgenden werden Verfahren und mögliche Ersatzmaße für Re-finding diskutiert. Es wird veranschaulicht, welche Konsequenz die Wahl eines Platzhalters für bestimmte Parameter wie *Anzahl der Nutzer*, *die Wiederfinden* und *Häufigkeit von Re-finding-Aktionen* besitzen.



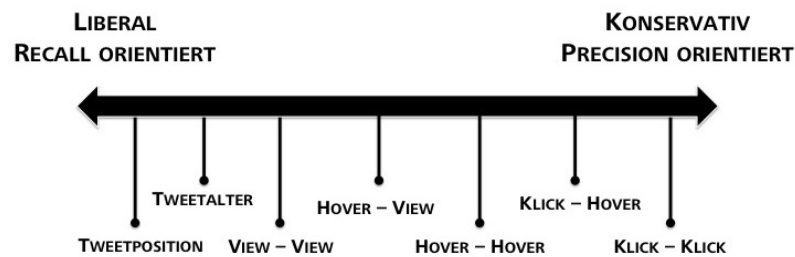
#### 6.4.4.1. Klassifikation von Re-finding-Sitzungen

Die Logdaten alleine liefern keine Hinweise darüber, ob der Nutzer die Intention besitzt, einen Tweet wiederzufinden. Aus diesem Zweck wurde das Erhebungsinstrument um das in Abschnitt 6.3.2.2 beschriebene *Experience Sampling* erweitert. Ziel des *Experience Sampling* war der Aufbau eines Goldstandards auf der Basis des qualitativen Feedbacks, mit dessen Hilfe ein Klassifikator trainiert werden kann, der die Sessions in Wiederfindenssitzungen und Sitzungen ohne Wiederfinden klassifiziert. Zur Erinnerung: Ein Experience-Sampling-Kurzfragebogen wurde dem Nutzer dann präsentiert, wenn er erneut auf einen Tweet klickt und die in 6.3.2.2 angegebenen Heuristiken zutreffen. Insgesamt wurden von den Probanden 96 Kurzfragebögen ausgefüllt, wobei 38,55% ( $n = 37$ ) Sessions als Wiederfindensaktionen gekennzeichnet werden. Anschließend wurde mit diesen Sessions und den Klasseninformationen ein Random-Forest-Klassifikator trainiert, der die restlichen Sessions in Re-finding bzw. Nicht-Re-finding-Sitzungen einteilt. Der Klassifikator charakterisiert über 25% der Sitzungen des Testdatensatzes (restliche Sessions) als Wiederfindenssitzungen, was als unrealistisch hoch einzuschätzen ist. Vergleicht man Re-finding in anderen Kontexten (Abschnitt 2.2.4), wäre wohl eine Zahl um die 10% als adäquat einzuschätzen. Eine mögliche Erklärung für den hohen Anteil an als Re-finding klassifizierten Sitzungen könnte sein, dass der Goldstandard nicht repräsentativ genug ist für das Klassifikationsproblem. Aus diesem Grund wurde davon abgesehen, den Goldstandard zur Klassifikation zu nutzen. Stattdessen musste ein Platzhaltermaß für die Klassifikation gefunden werden. Mögliche Platzhaltermaße werden im Folgenden diskutiert.

#### 6.4.4.2. Proxy-Diskussion: Häufigkeit des Auftretens von Re-finding-Verhalten

Neben einer Klassifikation von Re-finding-Sessions auf der Basis des Experience-Sampling-Goldstandards können Re-finding-Aktionen auch auf der Grundlage von *Proxies* (Platzhaltern) und dahinterstehenden Annahmen identifiziert werden. *Proxies* für Re-finding können in einem Kontinuum dargestellt werden und unterscheiden sich im Wesentlichen darin, ob sie eher liberal ausgerichtet sind und auf *Recall* (Vollständigkeit) abzielen, um die Gesamtheit aller Re-finding-Aktionen erfassen zu können, oder ob sie eher konservativ ausgerichtet sind und auf *Precision* (Genauigkeit) fokussieren. Ein recall-orientierter Proxy, der darauf abzielt, die Gesamtheit aller Wiederfindensaktionen zu erfassen, besitzt das Risiko einer hohen False-positive-Rate, also einer möglicherweise hohen Zahl an fälschlicherweise als Wiederfinden klassifizierten Sessions. Ein precision-orientierter Proxy versucht sicher zu gehen, dass die Sitzungen, welche durch den Proxy als Re-finding charakterisiert werden, auch tatsächlich Wiederfindenssitzungen sind, wobei die Gefahr einer höheren False-negative-Rate besteht, also dass Wiederfindenssitzungen als normale Sitzungen charakterisiert werden, obwohl sie Wiederfindensaktionen enthalten. Abbildung 6.19 verortet ausgewählte *Proxies* auf diesem eben beschriebenen Kontinuum.

Betrachtet man Tabelle 6.3 wird deutlich, dass nur etwa 30% aller Tweets eine Interaktion erfahren, was bedeutet, dass viele Tweets nicht gelesen werden, oder — was wahrscheinlicher erscheint — für die Rezeption eines Tweets keine Interaktion (in Form eines Hover oder eines Klicks) nötig ist. Vor diesem Hintergrund kann argumentiert werden, dass auch Tweeigenschaften wie z. B. die Tweetposition oder das Alter eines Tweets als *Proxies* genutzt werden können. Häufiger werden aber Interaktionssequenzen als *Proxies* instrumentalisiert, wobei die erneute Interaktion vorzugsweise in einer neuen Sitzung stattzufinden hat. Im Prinzip gibt es keine Grenzen für die Wahl eines *Proxies*. Auch Kombinationsmöglichkeiten zwischen Tweetfeatures und Interaktionssequenzen sind denkbar. Auszuschließen wären lediglich solche *Proxies*, deren dahinterstehende Annahmen unplausibel erscheinen. Der Proxy View — View in Abbildung 6.19 geht von der Annahme aus, dass ein Tweet wiedergefunden wird, wenn er in einer anderen Session erneut auf



**Abbildung 6.19.:** Platzhaltermaße (*Proxies*) verortet auf einem Kontinuum ihrer Ausrichtung von sehr liberal bis sehr konservativ

dem Bildschirm des Studienteilnehmers erscheint. Dies ist, wie der Abbildung 6.19 zu entnehmen ist, ein eher liberaler Proxy, der vermutlich einige Tweets als wiedergefunden interpretieren würde, für die dies nicht der Fall ist (*false positives*). Die hinter diesem Proxy stehende Annahme ist nicht unplausibel, da bei einer Social-Media-Plattform, die auf Echtzeitinformation ausgelegt ist, das erneute Auftreten eines Tweets auf dem Bildschirm als unwahrscheinlicher zu erachten ist, als beispielsweise das erneute Auftreten einer E-Mail in einem E-Mail-Client. Letztendlich hängt dies von mehreren Faktoren ab, die die Aktualität des Streams des Nutzers beeinflussen. Diese sind: die Nutzungsintensität des Studienteilnehmers, die Anzahl der Accounts, die der Teilnehmer folgt sowie der Tweetsendefrequenz dieser Accounts. All dies stellen Einflussfaktoren für diesen Proxy dar. Im Folgenden werden die Ersatzmaße Tweetalter, erneute Interaktion (*Hover — Hover*) sowie erneuter Klick (*Klick — Klick*) näher vorgestellt. Es wird erläutert, welche Konsequenz ihre Wahl für bestimmte Parameter wie *Vorkommenshäufigkeit von Re-finding*, *Anzahl an Re-finding-Sessions*, *Anzahl der wiedergefundenen Tweets* sowie *Anzahl an wiederfindenden Nutzern* besitzen.

#### **Tweetalter:**

Klassifiziert man Tweets, die zum Betrachtungszeitpunkt älter als der Median<sup>28</sup> des Tweetalters (1,4 Std. siehe Abschnitt 6.4.3.3) sind als Tweets, die wiedergefunden werden, resultiert dies in der Charakterisierung von 25,3% der gesehenen Tweets als wiedergefundene Tweets. In knapp 2500 Sessions würde Re-finding stattfinden, wobei beinahe alle (93,18%) Studienteilnehmer in Wiederfindensaktionen involviert wären. Die Zahlen deuten darauf hin, dass dies ein äußerst liberaler Proxy ist, der die Auftretenshäufigkeit von Wiederfinden überschätzt. Sein Einsatz würde in zahlreichen *false positives* resultieren. Einen genaueren Proxy stellt die erneute Interaktion — in diesem Fall das Überfahren mit der Maus (*Hover*) — in einer neuen Session dar (*Hover — Hover*). Mit 11% der Tweets, mit denen interagiert wurde, wurde auch in einer weiteren Session interagiert. Diese sind auf der Basis dieses *Proxies* als wiedergefundene Tweets zu interpretieren. Cross-Session-Interaktion führt dazu, dass 34 der 44 Probanden (77%) Tweets wiederfinden, wobei der Anteil an Tweets, die wiedergefunden werden zwischen 0,4% und 30% schwankt. Im Durchschnitt werden bei jedem Nutzer 10,1% aller Tweets wiedergefunden. Vermutlich überschätzt auch der Proxy *Hover — Hover* die Auftretenshäufigkeit von Re-finding, da auch eine erneute Interaktion in einer anderen Session stark von weiteren Einflussfaktoren abhängig ist.

Einen weiteren Proxy stellt der erneute Klick über eine Session hinweg dar, der, wie oben bereits erwähnt, auch häufiger in anderen Studien als Re-finding-Proxy verwendet wird (Teevan, Adar et al., 2007; Tyler & Teevan, 2010; Harvey & Elweiler, 2012). Im Kontext von Twitter ist dies ein sehr konservativer Proxy, da Tweets ohnehin sehr selten geklickt werden (nur 5% aller von Studienteil-

<sup>28</sup>Der Median ist hier letztlich auch nur ein arbiträr gewählter Wert. Es könnte auch der Mittelwert oder das 25%-Quantil sein. Letztendlich kommt es auch hier auf die Plausibilität des jeweiligen Wertes an.

Eindeutige Tweets für Nutzer 146735	Tweets interagiert	Tweets interagiert mit SI	Erneute Interaktion Tweets ohne SI	Erneute Interaktion Tweets mit SI
Hover oder Klick	40486 (27.59%)	4847 (11.97%)	375 (7.73%)	3996 (11.21%)
Nur Klick	8320 (5.67%)	4743 (57.00%)	134 (3.75%)	85 (1.79%)

**Tabelle 6.3.:** Vergleich der Häufigkeit, mit der Tweets mit und ohne spezieller Interaktion eine erneute Interaktion erfahren

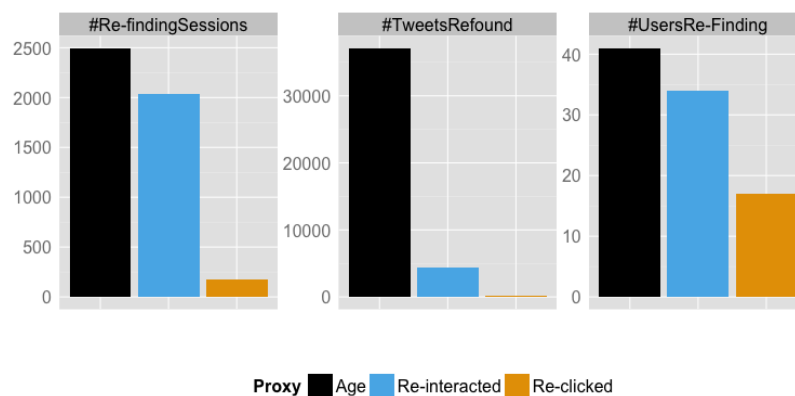
nehmern gesehenen Tweets wurden geklickt vgl. Tabelle 6.3) und es relativ unwahrscheinlich ist, dass Nutzer zufällig auf einen älteren Tweet klicken, auf den sie bereits in einer früheren Session geklickt haben, setzt man ein normales Nutzungsverhalten voraus. Verwendet man einen erneuten Klick (Re-Klick) in einer anderen Session als Proxy, resultiert dies in 276 Re-finding-Aktionen mit 219 einzelnen Tweets in 176 Sessions. Dies bedeutet zum einen, dass einzelne Sessions mehrere Re-finding-Aktivitäten enthalten können und zum anderen, dass 55 Tweets nicht nur einmal erneut geklickt werden, sondern in mehreren Sessions wiedergefunden werden können ( $Median = 2$ ,  $Max = 3$ ). 17 Studienteilnehmer sind in Re-finding-Aktionen involviert und zwischen 1,3% und 15,4% der geklickten Tweets werden von den Probanden wiedergefunden.

Abbildung 6.20 fasst die hier vorgestellten *Proxies* visuell zusammen und zeigt auf, dass die Wahl eines erneuten Klicks in einer weiteren Session zur deutlichen Abnahme aller hier diskutierten Wiederfindenscharakteristika führt. Der erneute Klick ist zweifellos der konservativste Proxy, welcher die Anzahl an *false positives* deutlich reduziert, aber dazu führt, dass nur 5% aller Tweets in Betracht gezogen werden und ein Großteil der Daten (95%), welche vermeintlich auch Re-finding-Instanzen enthalten können, außer Acht lässt. Obwohl das wahre Ausmaß des Re-finding-Verhaltens vermutlich in der Mitte des Kontinuums anzusetzen ist, wird für die weiteren Analysen der Proxy *Klick* — *Klick* genutzt, welcher auf *Precision* fokussiert, so dass bei der Darstellung der Ergebnisse davon ausgegangen werden kann, dass auch tatsächlich Re-finding-Verhalten beschrieben wird.

#### 6.4.4.3. Der nutzerbezogene Lebenszyklus eines Tweets

Unter dem nutzerbezogenen Lebenszyklus eines Tweets ist der Zeitrahmen zu verstehen, indem ein einzelner Tweet von Relevanz bzw. Interesse für einen Nutzer sein kann. Der zeitliche Rahmen ist dabei charakterisiert durch die Zeit, die zwischen dem ersten, initialen Lesen — bestimmt durch die Ersatzhandlung des Klicks auf einen der vielen möglichen Bereiche eines Tweets — und einem erneuten Klick, also dem Wiederfinden des Tweets, vergeht.

Der Boxplot in Abbildung 6.21 verdeutlicht, dass der nutzerbezogene Lebenszyklus eines Tweets länger ist, als man vermuten würde. Der Median liegt bei 18,8 Std., so dass die Hälfte aller Re-finding-Aktionen innerhalb eines Tages stattfinden. Im Durchschnitt werden Tweets aber erst nach 217 Std. wieder geklickt. Die Verschiebung ist durch das hohe Maximum zu erklären, welches zeigt, dass der längste Lebenszyklus eines Tweets in den Daten 1243,8 Std. bzw. 52 Tage beträgt. Tweets können also noch Wochen bzw. Monate nach ihrer ersten Betrachtung für den Nutzer von Relevanz sein. Dies ist abermals ein Hinweis auf Twitters informationsbewahrenden Charakter, der weit über die Perspektive von Twitter als Echtzeitinformationsnetzwerk hinausgeht. Es wird ebenfalls deutlich, dass die Zeit zwischen einem Tag und einer Woche von weniger erneuten Klicks gekennzeichnet ist. Erst nach etwa einer Woche beginnen wieder vermehrt Re-finding-Aktionen stattzufinden. Abbildung 6.21 verdeutlicht, wie Twitter-Funktionen (FAV/RT) und die zeitliche Dimension des Wiederfindens in Zusammenhang stehen. Teilt man die Daten beim Median in jeweils zwei Hälften — kurze Wiederfindensperiode ( $< Median = 18,8$  Std.) und lange Wiederfindensperiode ( $> Median = 18,8$  Std.) — ist zu erkennen, dass der Anteil der Tweets, welche eine spezielle

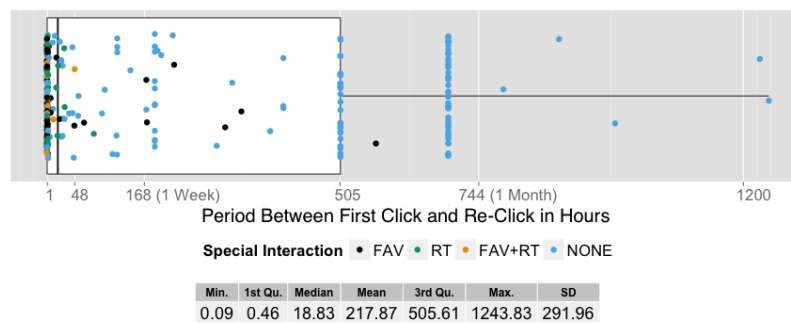


**Abbildung 6.20.:** Die drei vorgestellten Proxies Tweetalter, *Hover* — *Hover* und *Klick* — *Klick* und deren Konsequenz für die Parameter: Anzahl Wiederfindenssitzungen, Anzahl wiedergefundener Tweets, Anzahl wiederfindender Nutzer

Interaktion erfahren, im Bereich der kurzen Wiederfindensperiode mit 61,6% deutlich höher ist, als im Bereich der langen Wiederfindensperiode, in welcher nur 10,4% der Tweets favorisiert oder weitergeleitet werden. Auf den ersten Blick scheint dies kontraintuitiv zu sein, da in der Literatur Favorisieren und Retweeten häufig als Zeichen für Relevanz gewertet werden, respektive beide Features dazu genutzt werden, um Tweets in einem separaten Bereich aufzubewahren, um sie später einfacher wiederfinden zu können (boyd et al., 2010). Folglich müsste der Anteil an Tweets, welche eine spezielle Interaktion erfahren haben, in der langen Wiederfindensperiode höher sein, wobei aber das Gegenteil beobachtet werden kann. Ein Chi-Quadrat-Test bestätigt diese Beziehung und zeigt, dass spezielle Interaktion eine sehr starke Auswirkung darauf hat, ob ein Tweet nach kurzer oder langer Periode wieder geklickt wird. Tabelle 6.4 zeigt die Vierfeldertafel des Tests sowie die Testergebnisse. Das *odds ratio* oder Quotenverhältnis zeigt, dass die Wahrscheinlichkeit, dass ein Tweet nach einer langen Periode wiedergefunden wird, für Tweets ohne spezielle Interaktion 14 mal höher ist, als für Tweets, die favorisiert oder weitergeleitet wurden.

Die Ergebnisse legen nahe, dass Nutzer nur bedingt dazu in der Lage sind, die Relevanz eines Tweets für zukünftige Informationsbedürfnisse zu erkennen. Es fällt ihnen schwer zu beurteilen, was zu einem späteren Zeitpunkt wichtig sein könnte. Dieses Problem wurde in Kapitel 2 zusammen mit dem Konzept des *personal anticipated information need* (PAIN) eingeführt. Zusätzlich liegt die Vermutung nahe, dass Tweets vor allem dann favorisiert oder weitergeleitet werden, wenn sie für den Nutzer aktuell von Relevanz sind bzw. der Nutzer deren Bedeutung sofort einzuordnen weiß.

Neben der speziellen Interaktion kann auch der Effekt von *@Mentions* — ergo Tweets, welche mit einer Adressierung beginnen und somit Teil einer Konversation sind (Variable *POC*) — auf die Wiederfindensperiode analysiert werden. Liegt der Anteil an Tweets, die Teil einer Konversation sind, in der kurzen Wiederfindensperiode noch bei 21%, so ist keiner der wiedergeklickten Tweets in der langen Periode Teil einer Konversation. Aufgrund des Nullwertes kann ein Zusammenhangsmaß der beiden Variablen nicht über den herkömmlichen Chi-Quadrat-Test bestimmt werden. Campbell zeigt in einer Simulationsstudie, dass bei kleinen Stichprobenumfängen respektive bei sehr niedrigen beobachteten Werten, der  $N - 1$  Chi-Quadrat-Test nach Egon Pearson deutlich genauere p-Werte liefert, als der Standard Chi-Quadrat-Test nach Karl Pearson (Campbell, 2007). Busing, Weaver und Dubois greifen Campbells Studie auf, unterstützen dessen Empfehlung nachhaltig



**Abbildung 6.21.:** Boxplot (a) und Tabelle (b) der Zeitperiode (Lebenszyklus) zwischen erstem und erneutem Klick

und zeigen wie der herkömmliche Chi-Quadrat-Test, wie er in den Statistikprogrammen SPSS, SAS und R implementiert ist, angepasst werden kann (Busing et al., 2016).<sup>29</sup> Tabelle 6.5 zeigt die entsprechende Vierfeldertafel sowie die Testergebnisse. Es lässt sich ein signifikanter Zusammenhang zwischen den beiden Variablen *Teil einer Konversation*(POC) und *Dauer der Wiederfindensperiode* erkennen. Auch das *odds ratio* muss aufgrund des Nullwertes in der Vierfeldertafel in abgewandelter Form berechnet werden. Parzen et al. schlagen aufbauend auf früheren Arbeiten von Haldane sowie Gart und Zweifel vor, zu jeder Zelle eine Konstante  $c$  hinzuzufügen. In Simulationsstudien hat sich gezeigt, dass ein Wert von  $c = 0,5$  „[...] the smallest first-order finite sample bias of any estimate [...]“ (Parzen et al., 2002, S.422) besitzt.

Wie im Falle der speziellen Interaktion zeigt sich auch für Tweets, die Teil einer Konversation sind, dass diese mit Wiederfindensaktionen in Zusammenhang stehen. Das *odds ratio* belegt, dass die Wahrscheinlichkeit, dass ein Tweet nach einer langen Periode wiedergefunden wird, für Tweets ohne Zugehörigkeit zu einer Konversation 74 mal höher ist, als für Tweets, die mit einer *@Mention* beginnen und somit Teil einer Konversation sind. Der hohe Wert des *odds ratio* ist intuitiv und erklärt sich durch die Tatsache, dass in den Tweets, die in der langen Wiederfindensperiode erneut geklickt wurden, keine Tweets zu finden sind, die Teil einer Konversation sind (Null). Es lässt sich feststellen, dass Re-finding und die Zugehörigkeit zu einer Konversation in Zusammenhang stehen. Zeitlich gesehen erfolgt das Wiederfinden innerhalb eines Tages (kurze Wiederfindensperiode < Median), wie dies auch schon bei Tweets der Fall war, die favorisiert oder weitergeleitet wurden.

#### 6.4.4.4. Charakterisierung des Re-finding-Verhaltens auf Sessionebene

Legt man einen erneuten Klick über eine Session hinweg als Ersatzmaß für Wiederfindensereignisse zu Grunde, ergeben sich, wie bereits erwähnt, 176 Sessions in denen ein Wiederfinden stattfindet respektive 4397 Sitzungen ohne Re-finding. Um die Leitfrage nach dem Wiederfindensverhalten beantworten zu können, ist im Folgenden nun der Frage nachzugehen, wie sich Re-finding-Sessions von Sitzungen ohne Wiederfindensaktionen unterscheiden.

Nutzersitzungen sind durch 58 numerische/metrische Variablen charakterisiert. Diese Variablen entstehen durch die Anzahl an Ereignissen pro Eventkategorie (z. B. TIMELINEEVENTCOUNT), sowie dem prozentualen Anteil der Eventkategorie an den Gesamtereignissen in einer Session (z. B. TIMELINEEVENTPERCENT). Darüber hinaus werden auch die Anzahl der gesehenen Tweets (TWEETSSEEN), die Anzahl der geklickten Tweets — sowohl absolut (CLICKCOUNTPERTYPE) als auch relativ (CLICKTWEETPERCENT) — oder das durchschnittliche Alter (AVGTWEETAGE), der in

<sup>29</sup>  $N - 1\chi^2 = Pearson\chi^2 * \frac{N-1}{N}$

	Tweets mit S.I	Tweets ohne S.I	Gesamt
Kurze Periode (< Median)	85	53	138
Lange Periode (> Median)	14	124	138
Gesamt	99	177	276
$\chi^2=77.18$ , $df=1$ , $p<0.05$ , $\phi=0.536$ , Odds-Ratio: 14			

**Tabelle 6.4.:** Verteilung von Tweets mit und ohne speziellen Interaktion über kurze und lange Wiederfindensperioden

einer Session betrachteten Tweets, zur Charakterisierung einer Session herangezogen. Eine vollständige Liste der Variablen mit Erläuterung findet sich in Tabelle 6.7. Für die Überprüfung, inwiefern sich Re-finding-Sitzungen von Sitzungen unterscheiden, in denen kein Wiederfinden stattfindet, muss ein geeignetes statistisches Verfahren gefunden werden. Zunächst muss die Verteilung der Daten geprüft werden. Dies kann einerseits visuell (durch Histogramme oder Quantil-Quantil-Plots) andererseits durch Verwendung des Shapiro-Wilk-Tests geleistet werden. Der Shapiro-Wilk-Test vergleicht die Stichprobenwerte einer Variable mit Werten, bei denen eine Normalverteilung vorliegt. Erhält man ein nicht-signifikantes Ergebnis (bei einem Signifikanzniveau von  $\alpha = 0,05$   $p > 0,05$ ), so kann davon ausgegangen werden, dass die Verteilung der Werte nicht von einer Normalverteilung abweicht, ergo die Werte normalverteilt sind. Erhält man mit  $p < 0,05$  ein signifikantes Ergebnis, lässt dies den Schluss zu, dass die Daten nicht normalverteilt sind. Nach Field ergeben sich bei großem Stichprobenumfang sehr leicht signifikante Werte, weshalb zusätzlich eine visuelle Überprüfung mittels Histogramm und Q-Q-Plot angebracht ist (Field & Miles, 2014, S.182).

Aufbauend auf den Ergebnissen des Shapiro-Wilk-Test kann für alle Variablen die Normalverteilungsannahme abgelehnt werden (vgl. Anhang B.1). Auch die visuelle Begutachtung der Variablen stützt diese Tatsache. Abbildung 6.22 zeigt exemplarisch für dieses Vorgehen das Histogramm und den Quantil-Quantil-Plot für die Variable `DURATION`. Sowohl beim Histogramm als auch im Q-Q-Plot sind deutlich eine extrem rechtsschiefe Verteilung zu erkennen.<sup>30</sup> Bei den anderen Variablen wurde analog verfahren. Auch dort kann stets eine rechtsschiefe mit einem sog. *heavy tail* oder *long tail* versehene Verteilung beobachtet werden. Da alle Variablen erwiesenermaßen nicht normalverteilt sind, wurde auf eine Varianzhomogenitätsprüfung verzichtet.

Aufgrund der Non-Normalität der Datenverteilung muss auf nicht-parametrische Verfahren zurückgegriffen werden, um Verhaltensunterschiede zwischen Re-finding- und Nicht-Re-finding-Sessions zu testen. Eine Möglichkeit bietet der Rangsummentest nach Wilcoxon (Bortz & Schuster, 2011, S.133). Hierbei wird jedoch ein weiteres Problem evident: die stark unterschiedlichen Stichprobengröße für Wiederfindenssitzungen ( $n = 176$ ) und Sitzungen in denen kein Wiederfinden stattfindet ( $n = 4397$ ). Obwohl der Wilcoxon-Rangsummentest für ungleiche Stichproben erwiesenermaßen geeignet ist, liegen keine Simulationsstudien darüber vor, ob dies auch bei deutlich unterschiedlichen Umfängen (25fache Größe) noch der Fall ist. Eine Lösung für dieses Problem wäre die Herangehensweise, aus der gesamten Stichprobe der normalen Session 176 Sessions zufällig zu ziehen, und diese mit den 176 Re-finding-Sitzungen zu vergleichen. Da bereits das Ausgangssample ein zufälliges Sample ist, würde eine erneute zufällige Stichprobe den Effekt verstärken, dass beim Vergleich der beiden Sessiongruppen die Unterschiede in den einzelnen 58 Features eher zufällig zu Stande kommen, als dass sie auch auf eine Grundgesamtheit aller Sessions zutreffen würden. Als Konsequenz wurde zur Überprüfung der signifikanten Unterschiede zwischen Gruppen das Ver-

<sup>30</sup>Zur Orientierung: bei der Normalverteilung würden sich die Punkte im Q-Q-Plot entlang der schwarzen Linie ausrichten. Dies verdeutlicht abermals, wie stark die Verteilung der Variable `DURATION` von einer Normalverteilung abweicht.

	Tweets POC	Tweets nicht POC	Gesamt
Kurze Periode (< Median)	29	109	138
Lange Periode (> Median)	0	138	138
Gesamt	29	247	276
$\chi^2=32,40$ , $df=1$ , $p<0.05$ , $\phi=0.34$ , Odds-Ratio: 74,63			

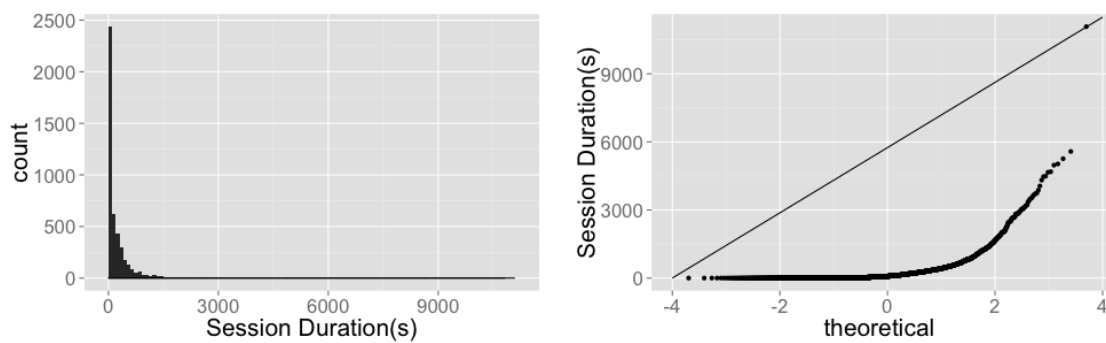
**Tabelle 6.5.:** Verteilung von Tweets mit und ohne @Mention über kurze und lange Wiederfindensperioden

fahren des nicht-parametrischen Bootstraps verwendet, wie bereits in Kapitel 5.7.1 beschrieben.

Es werden pro Gruppe  $B = 10000$  Stichproben der Größe  $n = 176$  gezogen, jeweils die Mittelwerte berechnet und voneinander subtrahiert. Schließlich entstehen 58 Grafiken, anhand derer abgelesen werden kann, ob der Nulleffekt existiert, also keine Unterschiede erkennbar sind, oder nicht. Diese sind dem Anhang B.2 zu entnehmen. Abbildung 6.23 zeigt exemplarisch die Variablen Sessiondauer (DURATION) und Anzahl an Suchereignissen (SEARCHEVENTCOUNT), die einen signifikanten, respektive nicht signifikanten Unterschied in den Mittelwert-Differenzen der Gruppen zeigen.

Insgesamt ergeben sich 28 Variablen, welche den Wert des Null-Effekts nicht aufweisen und somit als signifikant unterschiedlich gelten können. Es wird deutlich, dass Sitzungen, in denen Wiederfinden stattfindet, im Durchschnitt 40,1 (DURATION:  $Mean = 40,1 \mid 95\%KI = 21,9; 58,7$ ) mehr Events pro Session aufweisen und Re-finding-Sitzungen durchschnittlich auch um 261,6 Sekunden (EVENTSTOTAL:  $Mean = 261,6 \mid 95\%KI = 149,7; 371,2$ ) länger dauern. Interessanterweise finden in Sitzungen ohne Re-finding aber mehr Ereignisse in kürzerer Zeit statt, wodurch die Variable SESSIONBUSYNESSINDEX (definiert als Events pro Sessiondauer) höher liegt, respektive sich aufgrund der Mittelwertdifferenz ein mit  $-0,5$  negativer Wert für diese Variable ergibt. Aufgrund der Tatsache, dass sich beim Wiederfinden tendenziell mehr Ereignisse ereignen und Sessions auch länger dauern, ist es nicht verwunderlich, dass auch bei den meisten Eventkategorien — sowohl in der reinen Anzahl als auch prozentual gesehen — Wiederfindenssitzungen höhere Durchschnittswerte aufweisen. Interessanterweise lassen sich aber für die Ereigniskategorien Suche (SEARCHEVENTCOUNT/PERCENT) keine Unterschiede feststellen (siehe Abbildung 6.23). Dies scheint eher kontraintuitiv zu sein, da man in Wiederfindenssitzungen mehr Suchereignisse erwarten würde. Hier lässt sich ein erster Rückschluss auf Re-finding-Verhalten ziehen, das offenbar nicht durch klassische Suche — in Form einer Suchanfrage (Query) — durchgeführt wird. Durch die generelle Mehrzahl an Events und die längere Sessiondauer, erklärt sich auch die Tatsache, dass in Wiederfindenssitzungen mehr Tweets auf dem Bildschirm des Nutzers erscheinen (TWEETSSEEN:  $Mean = 39 \mid 95\%KI = 22,5; 55,6$ ), mehr Tweets durch Scrolling zur Timeline hinzugefügt werden (BUFFERCOUNTPERTYPE:  $Mean = 35,9 \mid 95\%KI = 19,7; 52,2$ ) mehr Tweets mit der Maus überfahren werden (HOVERCOUNTPERTYPE:  $Mean = 6,7 \mid 95\%KI = 2,39; 10,87$ ) und mehr Tweets geklickt werden, sowohl absolut (CLICKCOUNTPERTYPE:  $Mean = 3,48 \mid 95\%KI = 2,03; 4,94$ ) als auch relativ (CLICKTWEETPERCENT:  $Mean = 4,79 \mid 95\%KI = 3,02; 6,63$ ). Einzig beim Prozentsatz an mit der Maus überfahrenen Tweets kehrt sich das Vorzeichen um, so dass in Sitzungen ohne Wiederfindensaktionen im Durchschnitt zwischen 5,35 und 11,83% mehr Tweets *gehovered* werden als in Re-finding-Sitzungen.

Ein entscheidender Unterschied besteht auch in der durchschnittlichen und maximalen Position eines Tweets bzw. dem maximalen Alters eines Tweets, wenn mit diesen interagiert wird. Die durchschnittliche Position eines Tweets ist bei Re-finding-Sitzungen im Durchschnitt zwischen 4,5 bis 22,9 Positionen höher als dies bei normalen Sitzungen der Fall ist (AVGTWEETPOSITION:  $Mean = 13,6 \mid 95\%KI = 4,5; 22,9$ ). Die maximale Position eines Tweets liegt in Re-finding-



**Abbildung 6.22.:** Exemplarische Darstellung von Histogramm und Quantil-Quantil-Plot für die Variable DURATION zur visuellen Prüfung auf Normalverteilung

Sitzungen im Durchschnitt um 29,1 Positionen über dem Maximum der Position bei Sitzungen in denen kein Re-finding stattfindet (MAXTWEETPOSITION:  $Mean = 29,12 \mid 95\%KI = 4,5; 22,9$ ). Wie bereits erwähnt, ist auch das maximale Alter von Tweets in Re-finding-Sitzungen im Durchschnitt deutlich höher als dies bei normalen Sitzungen der Fall ist (MAXTWEETAGE:  $Mean = 6673364Sek. \mid 95\%KI = 1734958; 11709540$ ). Im Durchschnitt ist der älteste Tweet in einer Sitzung in Re-finding-Sessions um 77 Tage älter als bei normalen Sitzungen. Um zu älteren Tweets zu gelangen bzw. Tweets zu erreichen, die weiter unten in der eigenen *Timeline* oder den gesendeten Tweets von anderen Nutzern positioniert sind, ist vermeintlich mehr Mausinteraktion nötig. Tatsächlich lassen sich aber bei dieser Gruppe an Variablen keine signifikanten Unterschiede ausmachen (z. B. SCROLLSUMMARYCOUNT, DOWNSCROLLTOTAL).

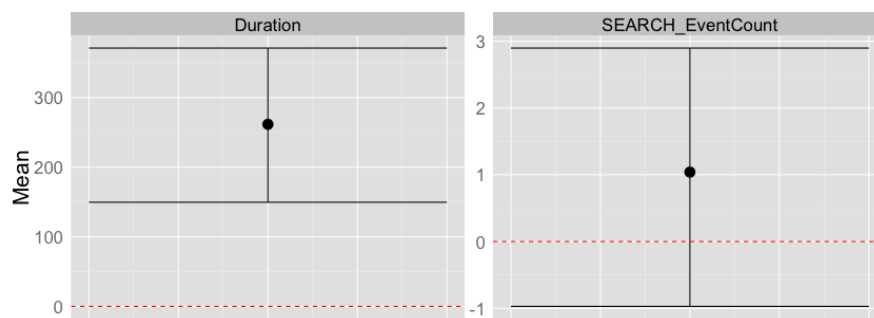
Ein weiterer Unterschied zwischen Re-finding und nicht Re-finding-Sitzungen lässt sich in der Anzahl der besuchten Nutzerprofile erkennen. In Re-finding-Sitzungen werden im Durchschnitt 0,4 mehr Profile besucht als in normalen Sitzungen (PROFILEVISITSPERSESSION:  $Mean = 0,4 \mid 95\%KI = 0,3; 0,8$ ). Auffällig ist, dass neben den eindeutigen Profilen (DISTINCTPROFILESVISITED:  $Mean = 0,25 \mid 95\%KI = 0,2; 0,5$ ) auch die *person uncertainty ratio* (PUR:  $Mean = 0,2 \mid 95\%KI = 0,04; 0,3$ ), also die Anzahl an erneut besuchten Profilen im Verhältnis zu den eindeutigen Profilen, die besucht wurden, bei Re-finding-Sessions signifikant höher ist als bei normalen Sessions. Auch wenn dieser Unterschied nur sehr gering ausfällt, deutet dies auf eine Redundanz oder Unsicherheit beim Besuch von Nutzerprofilen hin. Interessanterweise zeigt die Dauer der Profilbesuche aber keine signifikanten Unterschiede.

In Bezug auf die Leitfrage, wodurch Wiederfindensverhalten auf Sessionebene charakterisiert ist, lässt sich folgende Aussage treffen: Es ergibt sich ein Bild, das Re-finding-Sitzungen — aufgrund der Eventanzahl, Sessiondauer sowie der Interaktion mit Tweets — als intensivere Sessions charakterisiert. Die Daten zeigen, dass Alter und Position von Tweets ebenso eine wichtige Rolle beim Wiederfinden spielen wie die Profileseiten von anderen Nutzern. Schlüsselwort-Suche spielt im Re-finding offenbar eine untergeordnete Rolle, ebenso überrascht die Tatsache, dass keine signifikanten Unterschiede in Mausbewegungen wie Scrolling festgestellt werden können.

#### 6.4.4.5. Analyse von Wiederfindensstrategien

Zur weiteren Analyse des Re-finding-Verhaltens wird in einem nächsten Schritt der Blick auf den Kontext (URL) gelegt, in dem Tweets einen erneuten Klick erfahren. Der Kontext steht dabei in engem Zusammenhang mit der Strategie, welche Nutzer einsetzen, um zu diesem zu gelangen.

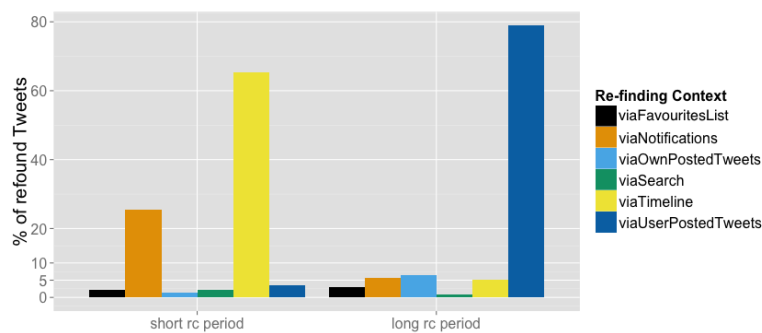




**Abbildung 6.23.:** Durchschnittliche Differenz der Mittelwerte der beiden Gruppen mit 95% Konfidenzintervall für die Variablen DURATION und SEARCHEVENTCOUNT. Die rotgestrichelte Linie zeigt den Nulleffekt an, der für einen signifikanten Unterschied nicht berührt werden dürfte.

Ein Beispiel: wird ein Tweet erfolgreich durch eine Suchanfrage wiedergefunden, wird er auf der SERP-Seite wiedergeklickt und somit im Kontext SEARCH. Um Strategieunterschiede bei unterschiedlichen Zeitperioden offenzulegen, können auch bei dieser Analyse die Daten anhand des Medians in kurze und lange Wiederfindensperioden eingeteilt werden. Abbildung 6.24 zeigt die Verteilung der Kontexte in kurze und lange Zeitintervalle seit dem ersten Klick. Es ist nicht überraschend, dass in kurzen Perioden der Kontext *Timeline* (65%) dominiert, da — solange sich die Anzahl neu eintreffender Tweets in Grenzen hält — durch einen vertretbaren Aufwand an Scrolling diese Tweets einfach im Stream lokalisiert werden können. Der zweithäufigste Kontext ist die NOTIFICATIONS-Seite, in welchem 25% aller Re-Klicks stattfinden. Strategien, die ein vorheriges Aufbewahren voraussetzen — wie das Wiederfinden über die eigenen gesendeten Tweets, wenn der Tweet vorher weitergeleitet (RT) wurde, oder das Lokalisieren über die Favoritenliste, wenn der Tweet vorher favorisiert wurde — kommen nur in 1,4% respektive 2,1% der Fälle vor. Auch die Suche wird nur in 2,1% der Wiederfindensfälle bemüht.

Betrachtet man die Wiederfindenskontexte der langen Wiederfindensperiode, zeigt sich ein deutlich anderes Bild. Das Wiederfinden über die eigene *Timeline* oder den Kontext NOTIFICATIONS sinkt auf Werte um die 5% ab. Re-finding über die Favoritenliste bleibt in etwa konstant bei 2,9%. Der Kontext SEARCH sinkt auf unter 1%. Die Strategie, einen Tweet über die eigenen gesendeten Tweets wiederzufinden, steigt und wird in etwa 5% der Fälle genutzt. Die dominanteste Wiederfindensstrategie ist jedoch das Lokalisieren des gesuchten Tweets in den gesendeten Tweets von anderen Nutzern. 78,9% aller Re-finding-Aktionen finden auf Profilseiten anderer Nutzer statt. Offensichtlich assoziieren die Studienteilnehmer Tweets mit den Tweetautoren bzw. Sendern eines Tweets und können sich gut daran erinnern — oder haben zumindest eine grobe Vorstellung davon — wer den Tweet gesendet hat. Dass zwischen Informationsobjekt, Autor und Rezipient eine starke Verbindung besteht, wurde bereits in anderen PIM-Studien mehrfach gezeigt. Aber auch in Bezug auf Twitter konnte die Eye-Tracking-Studie von Counts und Fisher zeigen, dass der Autor starken Einfluss auf das Erinnern eines Tweets besitzt: „factors like type of author impact memory for content (tweets from friends were remembered better than those from organizations)“ (Counts & Fisher, 2011, S. 11). Die Re-finding-Strategien der Studienteilnehmer ändern sich selten und bleiben zumindest in Bezug auf das Informationsobjekt konstant. Betrachtet man die Wiederfindenskontexte der 55 Tweets, die mehrmals wiedergefunden werden, findet in nur 10,10% (6 von 55) der Fällen ein Strategiewechsel statt, ergo der Tweet wird in einem anderen Kontext wieder-



**Abbildung 6.24.:** Vergleich der Re-finding-Strategien bei kurzen und langen Wiederfindensperioden. Als Proxy für die verwendete Strategie dient die jeweilige Twitter Unterseite.

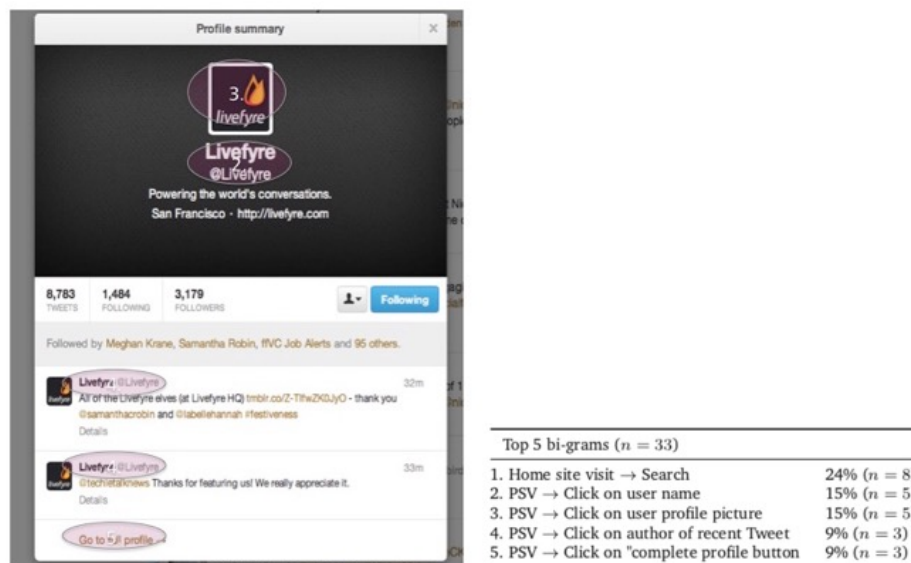
geklickt als beim ersten Wiederfinden. Auffällig dabei ist, dass bei einem erneuten Zugriff in 50% der Fälle Strategien zum Einsatz kommen, die auf vorheriger spezieller Interaktion, respektive *keeping* beruhen.

Twitter-Funktionen, welche in 6.4.4.3 als Mittel zur Verlängerung des Lebenszyklus eines Tweets identifiziert wurden, werden offenbar nur selten dazu genutzt, um Tweets wiederzufinden. Es bestehen damit mehrfach Diskrepanzen: Zum einen zwischen dem, was Nutzer vorgeben für PIM-Zwecke einzusetzen und ihrem tatsächlichen Verhalten. Zum anderen sind auch die Verhaltensweisen inkonsistent. Wenn Studienteilnehmer tatsächlich dazu in der Lage sind, die zukünftige Relevanz eines Tweets vorherzusagen und ihn deswegen korrekterweise favorisieren oder weiterleiten, werden diese Arten von Lesezeichen nicht oder nur selten dazu genutzt, um diese Tweets wiederzufinden. Das Nutzer bei einem ersten Kontakt mit einem Informationsobjekt dieses durch eine Art Bookmark in einen separaten Speicherplatz legen, das spätere Wiederfinden aber diesen Speicherplatz übergeht und auf eine andere Art und Weise stattfindet, wird in der PIM-Literatur als *Out-of-sight-out-of-mind-Problem* bezeichnet (Jones, Phuwanartnurak et al., 2005). Auch eine zeitversetzte Bewertung — unterstellt man den Probanden, dass diese Intention beim erneuten Besuch besteht — findet nicht statt. Die besondere Rolle des Autors bei Verhaltensweisen, die zum *Personal Information Management* auf Twitter gerechnet werden können, werden noch häufiger in der Arbeit angesprochen werden.

#### 6.4.4.6. Analyse von Re-finding-Pfaden

Da das Wiederfinden von Tweets über Profilseiten von anderen Nutzern (Kategorie: *USERPROFILEINT*) für lange Wiederfindensperioden als gängige Methode ausgemacht werden konnte, wird dieses Re-finding-Verhalten näher untersucht. Im Folgenden wird Wiederfinden auf unterster Ebene, der Event-Ebene, charakterisiert, indem analysiert wird, welche Pfade oder Wege Nutzer beschreiten, um zu den Profilseiten von anderen Nutzern zu gelangen. Inspiriert durch die Analysen von Wakeling, Clough und Sehn werden hierfür die Events als Bi-Gramme betrachtet, d.h. zwei konsekutive Events vor dem Besuch eines Nutzerprofils (*userProfileVisit*) bilden ein Bi-Gramm (Wakeling et al., 2014).

Die Tabelle 6.25 zeigt die Top 5 Bi-Gramm-Ereignisse, die vor einem Nutzerprofilbesuch auftreten. Das am häufigsten auftretende Bi-Gramm ist der Besuch einer eigenen Seite, gefolgt von einer Suchanfrage (24%). Insgesamt betrachtet entfallen 30% aller Ereignisse vor einem Nutzerprofil-

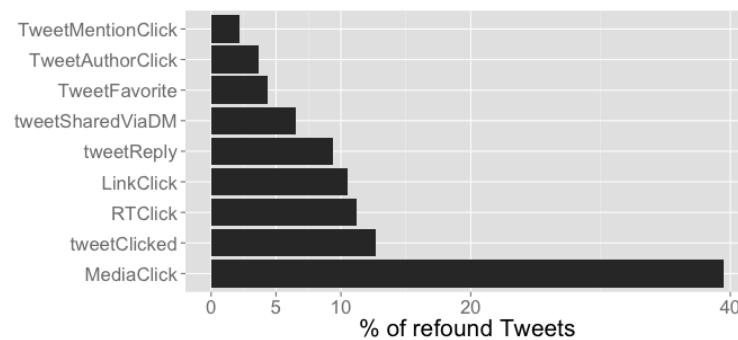


**Abbildung 6.25.:** Nutzerprofilzusammenfassung für den Account @Livefyre. Die beteiligten Links der in der Tabelle aufgeführten Bi-Gramme (2.-5.) sind farblich hervorgehoben. Top 5 Bi-Gramm-Ereignisse, die vor einem Nutzerprofilbesuch auftreten. PSV steht für *profile summary viewed* und ist das HTML-Overlay, welches die Zusammenfassung des Nutzerprofils zeigt.

besuch auf eine Suchanfrage. Dies deutet darauf hin, dass auch, wenn die direkte Suche nach einem Tweet — was in der Literatur als *teleporting* bezeichnet wird — eine eher untergeordnete Rolle spielt, Studienteilnehmer nach Personen suchen, um auf deren Nutzerprofil zu gelangen und dort per Navigations-Verhalten bzw. *orienting* den Tweet zu lokalisieren. Die Kombination aus Schlüsselwortsuche und anschließender Navigation charakterisiert Jones als das Abspringen über dem richtigen Informationsraum („parachute into the right region of an information space“ (Jones, 2013, S.29)), um anschließend via Navigation, die konkrete Information ausfindig zu machen. Die weiteren Bi-Gramme können auch in eine Gruppe klassifiziert werden und stellen ebenfalls eine Art Orienting-Verhalten dar. 70% der Bi-Gramme sind dadurch charakterisiert, dass nach einem Klick auf einen Tweetautor oder eine @Mention, die Profilzusammenfassung (*profileSummaryView*) des jeweiligen Nutzers erscheint. Von dort wird dann durch einen weiteren Klick das Nutzerprofil des jeweiligen Tweetautors besucht, wo schließlich ebenfalls mit Hilfe von Navigations-Verhalten der gesuchte Tweet wiedergefunden wird. In Abbildung 6.25 sind die jeweiligen Stellen der Profilzusammenfassung hervorgehoben, die mehrheitlich zur Weiternavigation genutzt wurden. Die Konzentration auf linkbasierte Navigation erinnert stark an die Navigation in Desktop-Dateisystemen und stellt klassisches Orienting-Verhalten dar. Die Analyse der Wiederfindenspfade mithilfe von Bi-Grammen zeigt auf, dass *teleporting* — also die direkte Suche nach einem Tweet — nicht auftritt, jedoch unterschiedliche Arten von Orienting-Verhaltensweisen an den Tag gelegt werden, um Tweets wiederzufinden. Beide Orienting-Verhaltensweisen unterstützen die in der Informationsverhaltensforschung geläufige These, dass Individuen als Quelle oder unterstützende Instanz bei der Suche nach Information bevorzugt werden. Darüber hinaus wird abermals die besondere Bedeutung des Tweetsenders bei Wiederfindensaktionen evident.

#### 6.4.4.7. Analyse von Wiederfindensgründen

Analog zu Abschnitt 6.4.4.5, in welchem der Kontext (URL) als Platzhalter für die Wiederfindensstrategie instrumentalisiert wird, kann die Art des Ereignisses (*event type*) beim erneuten Klick auf einen Tweet als Proxy für den Wiederfindensgrund gewertet werden. Da, wie in Abschnitt 6.3.1 bereits beschrieben, eine sehr detaillierte Erfassung der einzelnen Events erfolgte, kann über die Art des erneuten Klicks, ein Rückschluss auf die vermeintliche Nutzerabsicht erfolgen.



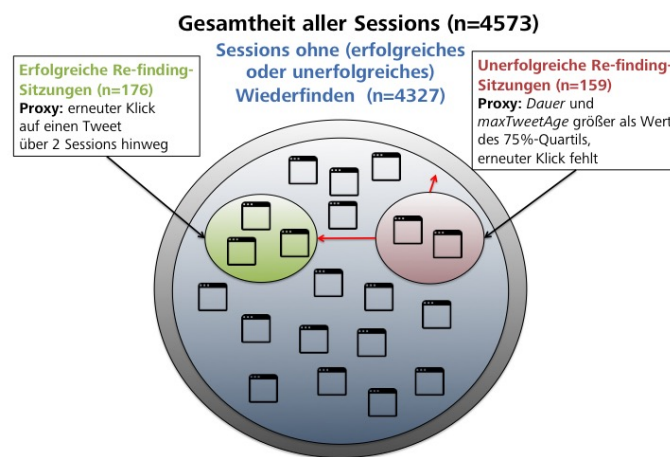
**Abbildung 6.26.:** Verteilung der Re-finding-Gründe für die 276 Re-finding-Aktionen

Abbildung 6.26 zeigt die Verteilung der Ereignisse, die zu den 276 Re-finding-Aktionen geführt haben und somit als Wiederfindensgründe gewertet werden können. Es zeigt sich, dass die Gründe sehr verschiedenartig sind, wobei mit knapp 40% der Großteil aller Tweets wiedergefunden wurde, um das Bild, Video oder die Website (*MediaClick*), die dem Tweet angehängt sind, erneut betrachten zu können. In 11,2% bzw. 4,3% der Fälle wurde ein Tweet wiedergefunden, um diesen weiterzuleiten (RT) bzw. zu favorisieren. Dies lässt vermuten, dass erst nach einer gewissen Beurteilungsphase der Nutzer die Entscheidung fällt, dass ein Tweet für seine Follower oder für ihn selber von Relevanz ist und dies durch entsprechende Interaktion vermittelt. Ein weiterer Erklärungsansatz könnte sein, dass das Wiederfinden sehr aufwändig und frustrierend war und der Tweet nun für ein erneutes Wiederfinden in einen separaten Speicher überführt wird. In 10,5% der Fälle wollen Nutzer den Link im Tweet wiederfinden, in etwa 10% der Fälle wird auf einen Tweet geantwortet. 6,5% der Wiederfindensaktionen sind motiviert durch das private Weiterleiten eines Tweets per Direktnachricht. Einen Tweet wiederzufinden, um darin erwähnte Personen (*TweetMentionClick*) oder den Autor wiederzufinden (*TweetAuthorClick*), ereignet sich nur in 3% bzw. 2% aller Gründe. 12,7% aller Re-Klicks entfallen auf das Ereignis *tweetClicked*, welches einen generischen Klick auf den Tweet darstellt, und nicht mit einer bestimmten Absicht in Verbindung zu bringen ist.

#### 6.4.4.8. Analyse von erfolglosem Re-finding-Verhalten

In den letzten Abschnitten lag der Fokus vor allem auf der Beschreibung erfolgreicher Wiederfindensinstanzen, wobei *erfolgreich* durch den Proxy *Re-klick* charakterisiert ist. Im Folgenden sollen vermeintlich erfolglose Re-finding-Sitzungen ausgemacht werden und abermals mit Hilfe der Bootstrapping-Methode mit erfolgreichen Re-finding-Sitzungen und normalen Twitter-Sessions

kontrastiert werden. Ziel ist es, Indikatoren zu finden, die auf Probleme bzw. Schwierigkeiten beim Wiederfinden hinweisen. Für *erfolgloses* Verhalten kann in diesem Fall auch nur ein Proxy genutzt werden, da ebenfalls keine qualitativen Daten bezüglich dieses Verhaltens vorliegen. Der Proxy für erfolgloses Re-finding beruht auf der Annahme, dass Sessions dann als erfolglos charakterisiert werden können, wenn sie Anzeichen für Re-finding-Verhalten zeigen, ergo Sessions mit hohen Werten für Dauer (DURATION) und Alter der betrachteten Tweets (MAXTWEETAGE) aufweisen, jedoch in diesen Sessions kein erneuter Klick auf einen Tweet verzeichnet werden kann. Beide Variablen wurden in Abschnitt 6.4.4.4 als Charakteristika für Re-finding-Sessions ausgemacht. Als Proxy werden jeweils die 75%-Quantil-Werte der beiden Variablen (DURATION 3.Quantil = 646 Sek.; MAXTWEETAGE 3.Quantil = 10990000 Sek.) verwendet und der Gesamtheit aller Sessions ohne erneuten Klick ( $n = 4397$ ) als Auswahlkriterien zu Grunde gelegt. Es ergeben sich 159 Sessions, die als erfolglose Wiederfindenssitzungen gewertet werden können. Abbildung 6.27 zeigt, wie sich die unterschiedlichen Sessiongruppen anhand der zugrundeliegenden Proxies zusammensetzen.



**Abbildung 6.27.:** Visualisierung der Zusammensetzung erfolgreicher, erfolgloser und normaler Sessions

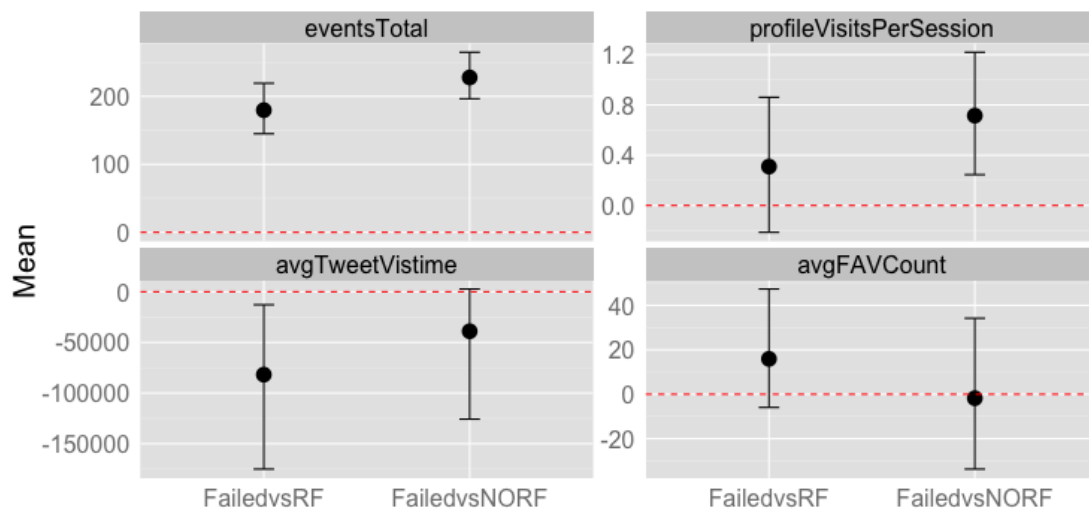
Die Annahmen, die hinter dem Proxy des erfolglosen Wiederfindens stehen, charakterisieren Wiederfinden als schwierige Aufgabe, die auch frustrierend sein kann. Dass die Schwierigkeit einer Aufgabe Einfluss auf das Verhalten des Nutzers hat, ist ein Umstand, der im Kontext von Suche mit Websuchmaschinen bereits häufig analysiert wurde (Arguello, 2014). In Abschnitt 6.4.4.4 konnte gezeigt werden, dass sich Nutzer in Wiederfindenssitzungen anders verhalten, als in normalen Sitzungen. Dabei wurden Sessionvariablen identifiziert, die charakteristisch für Wiederfindensverhalten sind und sich in einer Vorhersageaufgabe als gute Features erweisen (vgl. Abschnitt 6.4.5). In diesem Abschnitt soll zusätzlich gezeigt werden, wie sich erfolgloses Wiederfinden sowohl von erfolgreichen Wiederfindenssitzungen als auch von normalem Twitter-Verhalten unterscheidet und Indikatoren für schwierige Twitter-Nutzung ausgemacht werden. Eine plausible Annahme hierbei ist, dass sich erfolglose und erfolgreiche Wiederfindenssitzungen weniger stark unterscheiden als erfolglose und normale Sitzungen, da sich erfolglose und erfolgreiche nur anhand des erneuten Klicks unterscheiden und in den restlichen Session-Variablen ähnlicher sein sollten.

Tatsächlich ist dies der Fall: Bei beinahe allen Variablen gibt es eine Tendenz, die zeigt, dass der Unterschied zwischen erfolglosen Re-finding-Sitzungen und normalen Sitzungen deutlich höher

ausfällt, als dies zwischen erfolglosen und erfolgreichen Re-finding-Sitzungen der Fall ist. Folglich wird die oben formulierte Annahme durch die Daten gestützt, was die Plausibilität der beiden *Proxies* und deren dahinterstehenden Annahmen stärkt.

In Hinsicht auf signifikante Unterschiede zwischen den drei Session-Kategorien (erfolgreiche (RF) / erfolglose (*failed*) Wiederfindenssitzungen und normale Sitzungen (NORF)) lassen sich vier Szenarien beschreiben:

1. Beide Vergleiche zeigen keine signifikanten Unterschiede: Die Features dieser Kategorien sind weder dazu geeignet, um Wiederfinden vorherzusagen noch als Indikatoren um intensive/schwierige Twitter-Aufgaben zu erkennen.
2. Die betreffende Variable zeigt einen signifikanten Unterschied bei erfolglosen und erfolgreichen Wiederfindenssitzungen, nicht aber bei erfolglosen Wiederfindenssitzungen und normalen Twitter-Sitzungen. Diese Features können dazu genutzt werden, um schwierige Aufgaben vorherzusagen, sind jedoch weniger dazu geeignet, um zwischen normalen und Wiederfindenssitzungen zu unterscheiden. Dies geht auf die inhärente Ähnlichkeit von erfolgreichen und erfolglosen Wiederfindenssitzungen zurück.
3. Variablen zeigen signifikante Unterschiede bei Wiederfindenssitzungen vs. Sitzungen ohne Wiederfinden. Sie können zur Unterscheidung von Re-finding/Nicht-Re-finding genutzt werden, nicht aber für die Vorhersage, ob eine Aufgabe schwierig ist oder nicht.
4. Sessionvariablen, die für beide Kategorien signifikante Unterschiede aufweisen, sind Variablen, die sowohl bei der Charakterisierung von schwierigen Aufgaben als auch bei der Unterscheidung von Wiederfindenssitzungen und normalen Sitzungen nützlich sein können.



**Abbildung 6.28.:** Die vier unterschiedlichen Ergebnisszenarien bei der Analyse erfolglosen Verhaltens veranschaulicht anhand von vier verschiedenen Variablen

Abbildung 6.28 zeigt exemplarisch anhand der vier Variablen `EVENTSTOTAL`, `AVGTWEETVISTIME`, `PROFILEVISITSPERSESSION`, `AVGFAVCOUNT` die vier möglichen Ergebnisvarianten (siehe Anhang

		unerfolgreiche RF-Sitzungen vs. normale Sitzungen	
		signifikant	nicht signifikant
		eventsTotal sessionBusynessIndex directMSG_EventCount Notifications_EventCount Timeline_EventCount Tweeting_EventCount TweetsSeen maxRTCount maxFAVCount avgTweetPosition maxTweetPosition Buffer_CountPerType Buffer_TweetPercent Click_CountPerType Hover_CountPerType Hover_TweetPercent SummaryScrollCount DownScrollTotal UpScrollTotal mouseMoveCount totalDistance	FFMRelations_EventPercent avgTweetVstime Click_TweetPercent totalProfileTimeRelative avgMouseActivityIndex
erfolgreiche RF-Sitzungen vs. unerfolgreiche RF-Sitzungen	signifikant		
	nicht signifikant	FFMRelations_EventCount OwnProfileInt_EventCount Search_EventCount SingleTweet_EventCount UserProfileInt_EventCount maxTweetVstime profileVisitsPerSession totalProfileTime distinctProfilesVisited PUR avgDistance	Favorites_EventCount Lists_EventCount Other_EventCount directMSG_EventPercent Favorites_EventPercent Lists_EventPercent MUR avgRTCount avgFAVCount avgTweetAge scrollIndex

**Tabelle 6.6.:** Verortung aller Sessionvariablen in die vier möglichen Ergebnisszenarien als Ergebnis des Bootstrapping-Vergleichs

B.3 für die restlichen Variablen). Ein signifikanter Unterschied ist immer dann zu vermerken, wenn der Wert des Nulleffekts (Berührung der roten Linie) nicht auftritt. Tabelle 6.6 verortet schließlich die Sessionvariablen basierend auf den Bootstrapping-Ergebnissen in die vier oben beschriebenen Ergebniskategorien.

Tabelle 6.6 identifiziert mehrere Variablen, die zeigen, dass sich selbst erfolglose und erfolgreiche Wiederfindenssitzungen — die eine sehr starke Ähnlichkeit aufweisen sollten — in einigen Features unterscheiden. Die Variablen aus den Kategorien 1. und 3. sind diejenigen, die als Indikatoren für Schwierigkeit in diesem Abschnitt herausgearbeitet werden. Es wird deutlich, dass erfolgloses Wiederfinden intensive Arbeit ist, die sich durch mehr Events, belebtere Sessions, stärkerer Interaktion und Mausbewegungen kennzeichnet. Fraglich ist dabei, wie diese Information für eine zusätzliche Unterstützung beim Wiederfindensprozess genutzt werden kann, respektive, ob in Sessions, die zusätzlich einen deutlichen Unterschied zu Wiederfindenssitzungen aufweisen, eine stärkere Frustration beim Nutzer zu verzeichnen ist. Diese Frage kann ohne weitere Experimente nicht beantwortet werden. Ferner stellt sich die Frage, wie Session-Charakteristika von erfolglosen RF-Sitzungen zu interpretieren sind, die einen signifikanten Unterschied zu normalen Sessions aufweisen. Es wird deutlich, dass hier eine besondere Aufgabe vorliegt, die zu gesondertem Verhalten führt. Welches Verhalten dies genau ist, kann nicht bestimmt werden. Es wird jedoch deutlich, dass es intensivere, arbeitsreichere Sitzungen sind.

#### 6.4.5. Vorhersage von Re-finding-Zielen

Viele Erkenntnisse aus den vorherigen Abschnitten deuten darauf hin, dass Nutzer Probleme haben, den zukünftigen Nutzen eines Tweets richtig einzuschätzen und auf adäquate Weise — z. B. durch Favorisieren oder Weiterleiten — darauf zu reagieren. In einem nächsten Schritt wird daher analysiert, inwiefern maschinelle Lernverfahren dazu genutzt werden können, um den zukünftigen Wert eines Tweets vorherzusagen und damit zu bestimmen, ob ein Tweet wiedergefunden

wird oder nicht. Darüber hinaus stellt sich die Frage, welche Merkmale von Tweets und Charakteristika des Nutzerverhaltens gute Features für eine hohe Genauigkeit bei der Vorhersage sind. Die für Vorhersageprobleme verwendeten Merkmale werden in Abschnitt 6.4.5.2 vorgestellt und dort näher erläutert. Prinzipiell lässt sich sagen, dass immer dann, wenn ein bedeutender Unterschied zwischen Merkmalen von wiedergefundenen und zufällig ausgewählten Tweets festgestellt werden kann, diese Merkmale auch gute Features im Klassifikationsverfahren darstellen. In Abschnitt 6.4.4.4 wurde genau dies bereits für Merkmale aus der Kategorie *Session-Merkmale* durchgeführt, indem gezeigt wurde, wie sich Wiederfindenssituationen von normalen Twitter-Sitzungen auf Sessionebene unterscheiden.

#### 6.4.5.1. Experimentelles Setup und *Random Forest* als Klassifikationsverfahren

Die Vorhersageexperimente sind als binäres Klassifikationsproblem charakterisiert, bei denen bestimmt werden soll, ob ein Tweet wiedergefunden wird oder nicht. Zum Training eines Klassifikators können die Merkmale (vgl. Abschnitt 6.4.5.2) der 217 erneut geklickten Tweets genutzt werden, welche somit die positive Klasse bilden. Als negative Klasse dienen 217 Tweets, die zufällig aus der Gruppe der 8320 geklickten Tweets (vgl. Tabelle 6.3) gezogen werden. Bei dem Samplingverfahren wird berücksichtigt, dass die negative Klasse keinen *bias* enthält, was bedeutet, dass keine Tweets von Nutzern berücksichtigt werden, die niemals erneut auf einen Tweet klicken, ergo wiederfinden. Somit wurde die Menge an 217 Tweets erstellt, indem zunächst 100 Tweets pro Nutzer gezogen wurden, die auch tatsächlich einen Tweet wiedergefunden haben. Aus dieser Menge wurden schließlich die finalen 217 Tweets der negativen Klasse gezogen, welche zusammen mit den 217 Tweets der positiven Klasse die 434 Tweets bilden, die für das Training und Testen der Modelle herangezogen werden. Insgesamt wurden 15 verschiedene Modelle in unterschiedlichen Komplexitätsstufen miteinander verglichen. Abschnitt 6.4.5.2 erläutert die unterschiedlichen Feature-Gruppen bzw. Merkmale, deren Kombination zu den 15 Modellen führen, im Detail. Bei jeder Iteration wurde eine Validierung des erstellten Modells mit Hilfe des 10-fold-cross-validation-Verfahrens durchgeführt. Als Klassifikator wurde das maschinelle Lernverfahren *Random Forest* verwendet. Konkret wurde die von Liaw und Wiener in R implementierte Variante des Random-Forest-Verfahrens nach Breiman verwendet (Liaw & Wiener, 2002). Der Grund für die Wahl von *Random Forest* sowie die grundlegende Funktionsweise dieses Verfahrens werden im Folgenden kurz erläutert.

##### Das maschinelle Lernverfahren *Random Forest*

Das Random-Forest-Verfahren ist ein maschinelles Lernverfahren, das zur Kategorie der Ensemble-Learning-Methoden gehört. Diese Methoden zeichnen sich dadurch aus, dass sie eine Vielzahl an Modellen trainieren, deren Ergebnisse schließlich aggregiert werden bzw. zu einem Konsens-Modell zusammengefasst werden, welches dann die konkrete Vorhersage liefert (James et al., 2013, S.303). Vergleichbare Methoden sind das *boosting* oder das *bagging*. Ihnen allen gemein ist die Tatsache, dass sie eine Vielzahl an Entscheidungsbäumen (*decision trees*) trainieren und diese dann kombinieren. Grundsätzlich bezeichnet der Begriff *Random Forest* nicht ein einziges Verfahren, sondern eine Familie von ähnlichen Herangehensweisen, die eine zusätzliche Ebene der Randomisierung in den Baumbildungsprozess integrieren (Breiman, 2001, S.10).

Baumbasierte Methoden können sowohl zur Regression als auch zur Klassifikation verwendet werden und besitzen den Vorteil, dass sie relativ einfach zu erstellen sind und eine gute Interpretierbarkeit der Ergebnisse bieten (Hastie et al., 2009, S.305). Entscheidungsbäume können anderen Personen leicht erklärt werden und spiegeln menschliches Entscheidungsverhalten besser wieder als z. B. Regressionen (James et al., 2013, S.315). Einzelne Bäume können visualisiert werden, wobei höhergelegene Knoten wichtigere Variablen für die Vorhersage darstellen und die Blätter



der Bäume zeigen Variablen mit konkreten Werten, die das Ergebnis eines maschinellen Lernverfahrens greifbar machen. Zudem lassen sich aus Entscheidungsbäumen direkt Regeln ableiten, die in Algorithmen überführt werden können (James et al., 2013, S.306). In der Praxis besitzen sie zudem den Vorteil, dass nominal-skalierte Daten nicht in Dummyvariablen überführt werden müssen. Entscheidungsbäume leiden jedoch häufig an sehr hoher Varianz und einer zu starken Anpassung an die Trainingsdaten was zu *overfitting* führt. Bei Testdatensätzen erreichen sie zudem eine zu geringe Klassifikationsgenauigkeit (James et al., 2013, S.307). Generell sind Entscheidungsbäume im Vergleich zu anderen Lernverfahren wie *Support Vector Machines* (SVM) in Bezug auf die Klassifikationsleistung nicht wettbewerbsfähig (James et al., 2013, S.303). Um Nachteile zu kompensieren, werden mehrere Bäume kombiniert, was zu deutlichen Verbesserungen in der Vorhersageleistung führt. Jedoch hat bei der Kombination mehrerer Bäume die Interpretierbarkeit zu leiden (James et al., 2013, S.318). Am Beispiel von *Random Forest* kann aber gezeigt werden, dass diese Modelle keineswegs Black-Box-Verfahren sind, was vielen maschinellen Lernverfahren häufig unterstellt wird.

Eine Möglichkeit, mehrere Bäume zu kombinieren, bietet das von Leo Breiman entwickelte Random-Forest-Verfahren (Breiman, 2001). Prinzipiell stellt das Random-Forest-Verfahren eine Weiterentwicklung des Bagging-Verfahrens dar, welches auf der Bootstrap-Methode (*bagging* = *bootstrap aggregation*) basiert. Das bereits in Abschnitt 6.4.4.4 vorgestellte Bootstrapping kann dazu genutzt werden, um die Varianz und damit das *overfitting* einer statistischen Lernmethode zu reduzieren. Mit Hilfe von Bootstrapping werden  $B$  unterschiedliche Datensätze aus dem Ursprungsdatensatz gebildet und entsprechend viele unterschiedliche Entscheidungsbäume trainiert, deren Ergebnis schließlich gemittelt wird. Bei einem Klassifikationsproblem werden für jeden Baum, die zum Training verwendeten Beobachtungen sowie welche Klasse der Baum für die Beobachtung vorhergesagt hat, festgehalten. Die finale Klassenvorhersage für eine Beobachtung erfolgt dann über eine Mehrheitsentscheidung (*majority vote*) (James et al., 2013, S.318). *Random Forest* wird schließlich um eine weitere Variante der Randomisierung ergänzt, die zu einer De-korrelation der Bäume führt. Wie beim Bagging-Verfahren werden bei *Random Forest* eine Vielzahl an Entscheidungsbäumen auf der Basis von Bootstrap-Trainingsstichproben trainiert (Breiman, 2001, S.11). Neu hinzu kommt, dass der Algorithmus bei der Bildung eines jeden Baumes nicht die Gesamtheit aller Prädiktoren  $p$  zur Auswahl hat, sondern die Split-Variable lediglich aus einer zufälligen Stichprobe gewählt werden kann (in der Regel beträgt die Stichprobengröße  $m \approx \sqrt{p}$  sie kann aber auch die Größe  $m = 1$  besitzen) (Breiman, 2001, S.11). Aufgrund der Tatsache, dass bei jeder Baumbildung unterschiedliche Prädiktoren beteiligt sind, werden weniger häufig identische Blätter und Äste gebildet, wodurch die Bäume untereinander verschiedenartiger sind und somit weniger korrelieren (James et al., 2013, S.320). Beim Random-Forest-Verfahren kann eine Feinabstimmung des Algorithmus auf zwei Arten geschehen. Zum einen durch die Anzahl der Bäume, die gebildet werden (`ntree`) und zum anderen durch die Anzahl an Prädiktoren, die bei jedem Split zufällig gezogen werden (`mtry`). Für letzteres bietet das R-Paket eine eigene `tuneRF`-Funktion zur Bestimmung des optimalen Wertes für `mtry`. In der Praxis hat sich aber herausgestellt, dass bereits die Standardwerte gute Ergebnisse liefern (Fernández-Delgado et al., 2014).

Durch die zusätzlichen Stufen der Randomisierung beim Random-Forest-Verfahren reduziert sich auch die Einfachheit der Interpretation im Vergleich zu herkömmlichen Entscheidungsbäumen. *Random Forest* sind jedoch alles andere als Black-Box-Verfahren. Mit Hilfe der *variable importance* kann gezeigt werden, wie wichtig eine Variable für die Klassifikationsleistung ist (Breiman, 2001, S.23ff.). *partial dependence plots* zeigen, wie sich die Zielvariable verändert, wenn der Prädiktor verändert wird. Der Effekt anderer Variablen wird dabei nicht ignoriert, sondern die Vorhersagen der Zielvariable geschehen auf Basis des durchschnittlichen Effekts der restlichen Variablen unter besonderer Berücksichtigung des Prädiktors, dessen Abhängigkeit betrachtet werden soll (Hastie et al., 2009, S. 369-371). *proximity plots* können dazu genutzt werden, um zu zeigen, welche Be-

obachtungen aus der Perspektive des Random-Forest-Verfahren Ähnlichkeit besitzen (Hastie et al., 2009, S.595). Breiman fasst die positiven Charakteristika von *Random Forest* wie folgt zusammen (Breiman, 2001, S.10):

- Geringe Fehlerrate bei Vorhersagen im Vergleich zu anderen Verfahren (Adaboost)
- Robustheit gegenüber Ausreißern und Rauschen
- Schneller als andere Verfahren (speziell *bagging* und *boosting*)
- Verfahren liefert nützliche Informationen z. B. Fehlerraten, Variableneinfluss etc.
- Einfaches Verfahren, das auch leicht parallelisiert werden kann

Zusammenfassend lassen sich folgenden Gründe festhalten, die für die Wahl von *Random Forest* sprechen. Das Verfahren ist robust, in den verschiedensten Anwendungskontexten erprobt und liefert auch mit geringem Aufwand, also ohne starkes Tuning, sehr gute Klassifikationsergebnisse (Hastie et al., 2009). Fernández-Delgado und Kollegen zeigen in einer Evaluation von 179 Klassifikatoren aus 17 Familien anhand von 121 Datensätzen, dass *Random Forest* die Familie an Klassifikatoren ist, die die höchste durchschnittliche Genauigkeit bei allen Datensätzen liefert (Fernández-Delgado et al., 2014). Im vorliegenden Fall wurde auch mit SVM-Lernern experimentiert, wobei in einem ersten Schritt keine bessere Klassifikationsleistung erkennbar war. Neben der Klassifikationsgenauigkeit lag der Fokus auch auf der Interpretierbarkeit des Modells. Ziel des Vorhersage-experiments war es vor allem auch zu sehen, welche Features bzw. Featuregruppen den meisten Einfluss auf das Wiederfinden von Tweets besitzen. Dies kann mit Hilfe der *variable importance* veranschaulicht werden. Im Folgenden werden die für das Experiment genutzten Features näher erläutert.

#### 6.4.5.2. Beschreibung der verwendeten Features

Für das Vorhersage-Experiment wurden insgesamt 125 einzelne Features erhoben, respektive durch Feature-Engineering konstruiert. Diese Features lassen sich in vier Kategorien einteilen. Die Kategorien sind:

- Tweet-Merkmale
- Merkmale zur Nutzer-Sender-Beziehung
- Session-Merkmale
- Nutzer-Merkmale

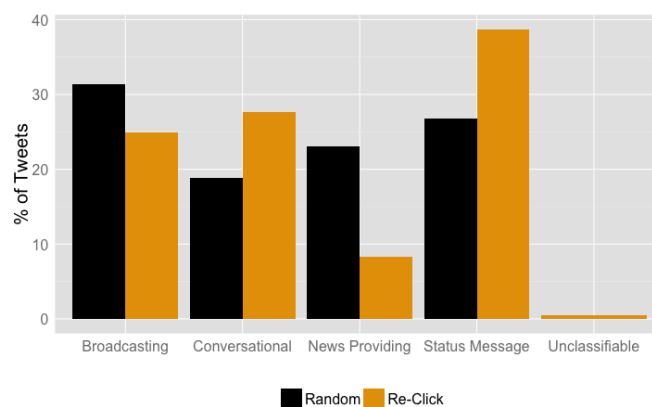
Tabelle 6.7 fasst die Merkmale tabellarisch zusammen und erläutert ihre Bedeutung.

**Tweet-Merkmale:** Zur Kategorie Tweet-Merkmale werden Features gerechnet, die allgemeine Tweet-Charakteristika beschreiben. Dies sind zum einen Merkmale in Bezug auf den Tweet-Text wie Anzahl der Wörter (WORDCOUNT), Anzahl der enthaltenen Zeichen (CHARCOUNT) oder die Part-Of-Speech-Informationen eines Tweets, wie die Anzahl an enthaltenen Hauptwörtern (21 Features zusammengefasst unter dem Featurenamen POSTAGS siehe Tabelle 6.7). Für das POS-Tagging des Tweet-Textes wurde ein spezielles, anhand von Tweets trainiertes Verfahren von Owoputi und Kollegen verwendet <sup>31</sup> (Owoputi et al., 2013). Zum anderen wurden tweetspezifische Merkmale festgehalten, wie die Anzahl an enthaltenen Hashtags und @Mentions (HASHTAGCOUNT und MENTIONCOUNT) oder FAV- und RT-Häufigkeit (FAVCount und RTCOUNT). Darüber hinaus wurde mit Hilfe von weiteren R-Paketen die Lesbarkeit (READABILITY.SCORE) sowie das Sentiment (SENTIMENT.SCORE) eines Tweets bestimmt. Die Lesbarkeit wurde mit Hilfe des *Flesch-Reading-Ease*

<sup>31</sup><http://www.cs.cmu.edu/ark/TweetNLP/>

errechnet.<sup>32</sup> Das Sentiment eines Tweets wurde mit Hilfe des R-Pakets `Sentimentr`<sup>33</sup> bestimmt, welches ein Verfahren zur Sentiment-Analyse nutzt, das wörterbuchbasiert arbeitet aber zusätzlich auch *valence shifters*, also Stimmungsänderungen, mit in die Beurteilung einfließen lässt.

Schließlich wurde die Kategorie eines Tweets mit Hilfe von zwei unbeteiligten Personen in einem Tagging- respektive Kodierverfahren bestimmt. Den Kodierern wurde ein Kodierschema vorgelegt, das anhand von Beispieltweets und Erklärungen die möglichen Tweetkategorien erläutert. Für die Tweetkategorien, dienten die Inhaltskategorien von Stephen Dann als Inspiration, welche für das Kodierverfahren etwas vereinfacht und angepasst wurden (Dann, 2010). Im Kodierverfahren fanden schließlich folgende sechs Inhaltskategorien Anwendung: *conversational*, *status message*, *news providing*, *broadcasting and pass along*, *unclassifiable and spam*. Die Bewerter durften mehrer Tag-kategorien vergeben, beschränkten sich im Allgemeinen aber auf eine Kategorie. Um die Güte des Verfahren zu beurteilen, wurde mit Hilfe des Cohens  $\kappa$  ein Maß der Urteilsübereinstimmung berechnet. Die Bewertung der beiden Kodierer wurde dann als übereinstimmend bewertet, wenn eine Übereinstimmung in zumindest einem Code pro Tweet vorhanden war. Die Bewerter konnten eine Übereinstimmung von  $\kappa = 0.48$  erreichen, was nach Grouven et al. einer *mittelmäßigen* Übereinstimmung entspricht (Grouven et al., 2007). Da für das Vorhersageexperiment eine finale Inhaltskategorie benötigt wird, wurde bei fehlender Übereinstimmung eine dritte Person zur Kodierung herangezogen und schließlich eine Mehrheitsentscheid gefällt. Das verwendete Kodierschema bzw. die unterschiedlichen Inhaltskategorien sind dem Anhang B.4 zu entnehmen.



**Abbildung 6.29.:** Aufteilung der Tweets in die unterschiedlichen Inhaltskategorien als Ergebnis des Kodierverfahrens

Abbildung 6.29 zeigt die Verteilung der Tweetkategorie für die wiedergefundenen und der zufällig für das Vorhersageexperiment ausgewählten Tweets. Die Verteilung wirkt eher kontraintuitiv, da der Großteil (38,7% und 27,7%) an wiedergefundenen Tweets einfache Status-message-Tweets bzw. Tweets, die Teil eines Dialogs (*conversational*) sind und somit durch eine *@Mention* zu Beginn des Tweets gekennzeichnet sind. Von wiedergefundenen Tweets würde man erwarten, dass diese interessante Nachrichten (*news providing*) enthalten. Der Anteil von Tweets aus der Kategorie *news providing* ist mit 8,3% innerhalb der wiedergefundenen Tweets gering ausgeprägt. Dass Tweets aus der Kategorie *conversational* häufig wiedergefunden werden, überrascht hingegen weniger. Bereits

<sup>32</sup><https://cran.r-project.org/web/packages/koRpus/index.html>

<sup>33</sup><https://github.com/trinker/sentimentr>

in Abschnitt 6.4.4.3 wurde ein spezieller Zusammenhang zwischen konversationellen Tweets und deren Lebenszyklus gezeigt. Allgemein betrachtet kann man erkennen, dass selbst das Wissen über die funktionale Kategorie eines Tweets nur bedingt eine Aussage über dessen Wiederfindenswahrscheinlichkeit zulässt.

Die Merkmale aus der Gruppe Tweet-Features können auch als Baseline für die Vorhersage bezeichnet werden, denn bis auf wenige Ausnahmen sind die meisten Merkmale auch via API auslesbar.

**Merkmale zur Nutzer-Sender-Beziehung:** Aus der Literatur ist bekannt, dass der Rezipient sich gut an den Autor respektive Sender eines Tweets erinnern kann (Counts & Fisher, 2011). Zudem konnte in den Abschnitten 6.4.4.5 und 6.4.4.6 die Bedeutung von Nutzerprofilen für den Prozess des Wiederfindens gezeigt werden. Mit den Merkmalen der Kategorie Nutzer-Sender-Beziehung wird der Bedeutung von Tweetautoren respektive -sendern Rechnung getragen. Die Kategorie übersetzt die Beziehung zwischen Nutzer und Sender in mehrere Features, wie die Anzahl an Tweets, die ein Nutzer von einem Sender erhält (TWEETSPERSENDER) und den Anteil an diesen Tweets, die der Nutzer klickt (TWEETSLICKEDPERSENDER), favorisiert oder weiterleitet (TWEETSSIPERSENDER). Darüber hinaus enthält die Kategorie Features mit Bezug zu grundlegenden Accountinformationen des Senders, wie Anzahl gesendeter Tweets (SENDERTWEETSPOSTED) oder Anzahl an Followern (SENDERFOLLOWERCOUNT) etc.. Zudem wird das Sendeverhalten des Senders berücksichtigt, indem der Anteil an weitergeleiteten Tweets (TWEETSRT) oder der Anteil an gesendeten Tweets, die Links enthalten (FRACTTWEETSLINKS) als Merkmale umgesetzt werden. Schließlich werden die Senderinformationen dazu genutzt, um eine heuristische Kategorisierung des Senders in die von Naaman et al. identifizierten Nutzerkategorien INFORMERS, also Nutzer die überwiegend informationellen Inhalt verbreiten und MEFORMERS, also Nutzer die überwiegend über sich selbst und ihren Alltag berichten, durchzuführen (Naaman et al., 2010).

**Session-Merkmale:** Session-Merkmale nehmen Bezug auf das Nutzerverhalten in der Session, in der ein Nutzer einen Tweet zum ersten Mal geklickt hat. Sie wurden bereits in Abschnitt 6.4.4.4 und 6.4.4.8 dazu genutzt, um Unterschiede zwischen normalen Sitzungen und Wiederfindenssitzungen festzustellen und Wiederfindensverhalten zu charakterisieren. Da dort bereits einige Charakteristika von Wiederfindenssitzungen evident wurden, ist es sehr wahrscheinlich, dass Merkmale, die dort als diskriminierende Features identifiziert wurden — wie Sessiondauer (DURATION) oder Anzahl an Events (EVENTSTOTAL) — sich diese auch für das Vorhersageexperiment als wichtige Features herausstellen werden. Die Merkmalskategorie beinhaltet *high-level* Sessioncharakteristika wie Sessiondauer, Anzahl an Events und Anzahl an Events pro Eventkategorie (TimelineEVENTS, USERPROFILEEVENTS). Darüber hinaus werden Merkmale in Bezug auf die in einer Session auftretenden Tweets betrachtet, sowohl in Bezug auf die Interaktion des Nutzers mit diesen, wie die Anzahl der geklickten Tweets (TWEETSLICK), als auch Tweetcharakteristika, wie beispielsweise die durchschnittliche und maximale Position der aufgetretenen Tweets (TWEETPOSITION). Das Merkmal MUR (Message Uncertainty Ratio) setzt die Anzahl der geklickten Tweets ins Verhältnis zu den eindeutigen Tweets, die geklickt wurden und versucht die Unsicherheit des Nutzers bei einer möglichen Wiederfindensaktion zu erfassen. Ferner werden die Mausbewegungsdaten innerhalb einer Session dazu genutzt, um eine Reihe an Merkmalen zu konstruieren, welche die Intensität der Mausbewegung und des Scrollings repräsentieren. Hierzu zählen die Anzahl an Pixel, die beim nach unten (DOWNSCROLLTOTAL) bzw. nach oben Scrollen (UPSCROLLTOTAL) zurückgelegt werden, oder auch die Gesamtfläche (AVGMOUSEACTIVITYINDEX), welche mit der Maus abgefahren wird. Schließlich wird auch bei den Session-Merkmalen die Bedeutung von Nutzerprofilen mit Merkmalen wie der Anzahl der besuchten Profile (PROFILEVISITS) oder der auf Profilen verbrachten Gesamtzeit (TOTALPROFILETIME) berücksichtigt. Ähnlich wie das Merkmal MUR versucht das Feature PUR die Unsicherheit in Bezug auf die besuchten Nutzerprofile zu erfassen.

**Nutzer-Merkmale:** Zur Kategorie Nutzer-Merkmale, zählt zum einen, das durch die Erweiterung

aufgezeichnete Twitter-Nutzungsverhalten im Allgemeinen, ergo die Gesamtzahl an Sessions eines Nutzers (SESSIONCOUNT), die Anzahl an Sessions pro Tag (SESSIONPERUSERPERDAY) oder die durchschnittliche Sessiondauer (MEANSESSIONDURATION). Zum anderen fließen die Accountinformationen des Nutzers, wie die Anzahl der geposteten Tweets (USERTWEETSPOSTED) oder die Anzahl an Favoriten (FAVOURITESCOUNT), als Merkmale in diese Kategorie ein.

Einige dieser Features lassen sich sehr leicht erheben. Die meisten Merkmale von Tweets (WORDCOUNT oder HASHTAGCOUNT) sind einfach auslesbar und sind auch über die API abgreifbar; das Wissen über einen Klick auf den jeweiligen Tweet vorausgesetzt. Andere Merkmale z. B. aus der Gruppe *Nutzer-Sender-Beziehung* erfordern die Erhebung und Speicherung der Nutzer-Sender-Interaktionsdaten über einen längeren Zeitraum hinweg. Wie bereits erwähnt, können die Merkmale aus der Gruppe Tweet-Merkmale als *baseline* für die Vorhersage erachtet werden, da der Großteil dieser Informationen auch über die API ausgelesen werden kann. Abschließend lässt sich sagen, dass die Anzahl und Art der verwendeten bzw. vorgestellten Merkmale keinesfalls erschöpft ist und durch Feature-Engineering noch weitere Merkmale aus den Daten extrahiert werden können.

#### 6.4.5.3. Ergebnisse des Vorhersage-Experiments

Tabelle 6.30a zeigt die Vorhersageergebnisse der 15 Modelle. Als Maße für die Güte der Vorhersageleistung werden das F1-Maß (als harmonisches Mittel aus Precision und Recall), die Genauigkeit sowie der AUC-Wert angegeben. Zusätzlich wird die Leistung der Klassifikatoren anhand von ROC-Kurven in Abbildung 6.30b visualisiert.

Sowohl anhand der Gütemaße als auch anhand der ROC-Kurven kann man erkennen, dass alle Modelle gute Vorhersageergebnisse liefern und in ihrer Qualität vergleichbar sind. ROC-Kurven sollten sich bei einem optimalen Klassifikationsergebnis der linken oberen Ecke annähern. In Abbildung 6.30b lässt sich dies für das Modell erkennen, welches mit allen Feature-Gruppen arbeitet und mit 0,88 den höchsten AUC-Wert besitzt, wobei viele weitere Modelle einen ähnlich hohen AUC-Wert besitzen. Das Baseline-Modell hebt sich auf negative Weise von den restlichen Modellen ab, da es in der Ergebnistabelle (vgl. Tabelle 6.30a) auf dem letzten Platz zu finden ist. Auch die dazugehörige ROC-Kurve (blaue Linie in Abbildung 6.30b) liegt einem rein zufälligen Klassifikator (Diagonale der ROC-Kurve) am nächsten. Dennoch ist es bemerkenswert, dass das Wissen, welcher Tweet geklickt wurde und dessen Charakteristika ausreichen, um mit einer Genauigkeit von 72% vorherzusagen, welcher Tweet in einer zukünftigen Sitzung wiedergefunden wird. Ebenfalls lässt sich das Muster, welches in Abschnitt 5.7.8 evident wurde, bestätigen: Nutzer, die Twitter häufiger und intensiver nutzen, sind auch diejenigen Nutzer, die häufiger wiederfinden. Dies erklärt, weshalb ein Modell nur auf der Basis von Nutzerinformationen besser abschneidet, als ein Modell, das Tweet-Features zur Klassifikation nutzt.

Das Modell mit der besten Klassifikationsleistung — betrachtet man das F1-Maß ( $F1 = 0,80$ ) — (vgl. Zeile 1 Tabelle 6.30a) kombiniert Merkmale aus den Feature-Gruppen: Tweet-Merkmale, Merkmale zur Nutzer-Sender-Beziehung und Session-Merkmale. Insgesamt besteht das Modell aus 115 Features, welche auch auf Langzeitbeobachtung des Nutzerverhaltens basieren.

In der Praxis<sup>34</sup> sind zwei Perspektiven auf die Kombination der Feature-Gruppen bzw. die Anzahl und Art der zur Klassifikation verwendeten Merkmale von Interesse. Zum einen ist dies der Aufwand, der für die Sammlung und Speicherung bestimmter Merkmale benötigt wird. In diesem Unterkapitel wurde bereits darauf hingewiesen, dass Merkmale zur Nutzer-Sender-Beziehung sowie Nutzer-Merkmale eine Langzeitbeobachtung des Nutzers erfordern und somit arbeitsintensiv

<sup>34</sup>Denkbar wäre ein Assistenzsystem, welches dem Nutzer bei potentiellen Wiederfindensaktionen Hilfestellung leistet. In Abschnitt 8 der Arbeit werden noch einige Designvorschläge vorgestellt, die zeigen, wie diese Hilfestellung aussehen kann.

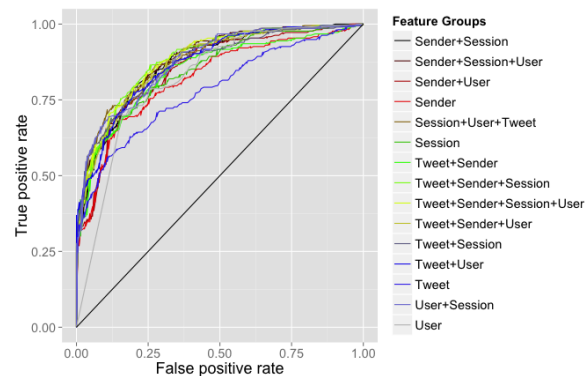
Name	Beschreibung
<i>Tweet-Features</i>	
TweetAge	Alter des Tweets zum Zeitpunkt des Klicks
TweetPosition	Position im Stream zum Zeitpunkt des Klicks
TweetVistime	Sichtbarkeit bis zum Zeitpunkt des Klicks
SpecialInteraction	Wurde der Tweet favorisiert (FAV) und/oder weitergeleitet (RT)?
FAVCount	Anzahl Favorisierungen zum Zeitpunkt des Klicks
RTCount	Anzahl RTs zum Zeitpunkt des Klicks
MentionCount	Anzahl der @Mentions in einem Tweet
LinkCount	Anzahl der Links in einem Tweet
HashtagCount	Anzahl der Hashtags in einem Tweet
maxFAVCount	Gesamtzahl der Favorisierungen eines Tweets
maxRTCount	Gesamtzahl der RTs eines Tweets
CharCount	Anzahl der Zeichen in einem Tweet
WordCount	Anzahl der Wörter in einem Tweet
avgWordLength	Durchschnittliche Wortlänge
POSTags	21 Part-of-Speech-Merkmale z. B. Anzahl Hauptwörter
POC	War der Tweet Teil einer Konversation?
sentiment.score	Positives oder negatives Sentiment eines Tweets
readability.score	Lesbarkeitswert nach Flesch
content.cat	Tweet-Kategorie ermittelt durch das Kodierungsverfahren
<i>Merkmale der Nutzer-Sender-Beziehung</i>	
TweetsPerSender	Anzahl der Tweets, die der Nutzer von Sender X erhalten hat
TweetsClickedPerSender†	Anzahl geklickter Tweets erhalten von Sender X
TweetsSIPerSender†	Anzahl Tweets mit spezieller Interaktion (FAV/RT) erhalten von Sender X
SpecialInteractionNorm	Normalisierter SI-Wert in Bezug auf generelles SI-Interaktionsverhalten
FracTweetsMentions	Anteil der Tweets von Sender X die @Mentions enthalten
FracTweetsHashtags	Anteil der Tweets von Sender X die Hashtags enthalten
FracTweetsLinks	Anteil der Tweets von Sender X die Links enthalten
TweetsRT†	Anzahl und Anteil an Tweets die von Sender X weitergeleitet wurden
SenderTweetsPosted	Anzahl an gesendeten Tweets von Sender X
SenderFollowerCount	Anzahl an Follower von Sender X
SenderFollowingCount	Anzahl Accounts denen Sender X folgt
SenderFAVCount	Anzahl an von Sender X favorisierten Tweets
SenderCreated	Zeit seit der Registrierung
SenderListedCount	Anzahl an Listen in denen Sender X gelistet ist
UserClass	Meformer oder Informer
<i>Session-Features</i>	
EventsTotal	Anzahl an Events pro Session
Duration	Dauer der Session in Sekunden
BusynessIndex	Anzahl an Events/Dauer der Session
TimelineEvents†	Anzahl an Events in der Eventkategorie Timeline
NotificationsEvents†	Anzahl an Events in der Eventkategorie Notifications
UserProfileEvent†	Anzahl an Events in der Eventkategorie USERPROFILEINT
OwnProfileEvent†	Anzahl an Events in der Eventkategorie OWNPROFILEINT
FFMRelationsEvents†	Anzahl an Events in der Eventkategorie FFMRELATIONS
SearchEvents†	Anzahl an Events in der Eventkategorie SEARCH
FavouritesEvents†	Anzahl an Events in der Eventkategorie FAVOURITESLIST
SingleTweetEvents†	Anzahl an Events in der Eventkategorie SINGLETWEET
TweetingEvents†	Anzahl an gesendeten Tweets in der Session
DirectmessageEvents†	Anzahl an Events in der Eventkategorie DIRECTMESSAGE
ListEvents†	Anzahl an Events auf Listen
OtherEvents†	Anzahl an Events auf weiteren Unterseiten (z. B. Twitter API)
TweetsSeen	Anzahl an auf dem Bildschirm erschienenen Tweets pro Session
MUR	Message Uncertainty Ratio (Geklickte Tweets/Eindeutige geklickte Tweets)
RTCountSession*	Durchschnitt und Maximum an RT für die in der Session betrachteten Tweets
FAVCountSession*	Durchschnitt und Maximum an FAV für die in der Session betrachteten Tweets
TweetPosition*	Durchschnitt und Maximum der Position für die in der Session betrachteten Tweets
TweetAge*	Durchschnitt und Maximum des Alters für die in der Session betrachteten Tweets
TweetVistime*	Durchschnitt und Maximum der Dauer bis zur Interaktion für die in der Session betrachteten Tweets
TweetsBuffer†	Tweets zum Stream hinzugefügt durch Scrolling
TweetsClick†	Tweets geklickt
TweetsHover†	Tweets gehovered
ScrollSummary	Anzahl an kontinuierlichen Scrolling-Aktivitäten
DownScrollTotal	Gesamtdistanz zurückgelegt durch Scrolling nach unten in Pixel
UpScrollTotal	Gesamtdistanz zurückgelegt durch Scrolling nach oben in Pixel
ScrollIndex	Gesamtzahl an Pixel/Gesamtzahl an Scrolling-Events
MouseMoveCount	Anzahl an einzelnen Mausebewegungen
MouseDistance*	Durchschnittliche und Maximum an Distanz in Pixel
AvgMouseMoveIndex	Fläche der Seite (in $Pixel^2$ ) abgefahren durch Mausebewegungen
ProfileVisits	Anzahl an besuchten Nutzerprofilen
TotalProfileTime	Dauer der Nutzerprofilbesuche in Sekunden
DistinctProfilesVisited	Anzahl an eindeutigen Nutzerprofilbesuchen
PUR	Person Uncertainty Ratio (Nutzerprofilbesuche/Eindeutige Nutzerprofilbesuche)
<i>Nutzer-Merkmale</i>	
MeanSessionDuration	Durchschnittliche Sessionlänge in Sekunden für Nutzer Y
SessionCount	Anzahl an Sessions für Nutzer Y
AvgEventPerSession	Durchschnittliche Anzahl an Events pro Session für Nutzer Y
SessionPerUserPerDay	Anzahl an Events pro Session pro Tag für Nutzer Y
UserTweetsPosted	Anzahl gesendeter Tweets von Nutzer Y
UserFollowerCount	Anzahl an Follower von Nutzer Y
UserFollowingCount	Anzahl an Accounts die Nutzer Y folgt
UserFAVCount	Anzahl an favorisierten Tweets von Nutzer Y
UserCreated	Zeit seit der Registrierung
UserMediaSent	Anzahl an gesendeten Tweets mit medialem Inhalt von Nutzer Y

**Tabelle 6.7.:** Überblick über die Features, die für die Vorhersage genutzt wurden, klassifiziert in Feature-Gruppen. Features, die mit einem † gekennzeichnet sind, haben absolute und relative Werte. Features, die mit einem \* versehen sind, besitzen einen Maximal- und einen Durchschnittswert.

sind in Bezug auf Sammlung und Speicherung dieser Daten. Zum anderen wäre eine Optimierung der Berechnung der Klassifikation durch den Random-Forest-Algorithmus eine Perspektive, die bei der Auswahl der verwendeten Merkmale berücksichtigt werden kann. Ideal wäre eine möglichst niedrige Anzahl an Variablen, die zu einer ebenso guten Klassifikationsleistung führt, wie dies bei der Nutzung des gesamten Feature-Raums möglich wäre, um eine möglichst niedrige Rechenkom-

	Feature-Gruppen	F1-Maß	Accuracy	AUC	Anzahl Features
1	Tweet + Sender + Session	0,80	0,79	0,88	114
2	Tweet + Session	0,79	0,78	0,87	97
3	Sender + Session	0,79	0,78	0,87	76
4	Tweet + Sender + User	0,79	0,79	0,87	67
5	Tweet + Sender + Session + User	0,79	0,78	0,88	125
6	Tweet + User	0,79	0,79	0,87	49
7	Sender + Session + User	0,78	0,78	0,88	86
8	Tweet + Session + User	0,78	0,77	0,88	107
9	Tweet + Sender	0,77	0,76	0,84	57
10	User + Session	0,77	0,77	0,88	68
11	Session	0,76	0,76	0,85	58
12	Sender + User	0,76	0,76	0,85	28
13	User	0,75	0,75	0,82	10
14	Sender	0,75	0,74	0,82	18
15	Tweet	0,74	0,72	0,78	39
	<b>Top-3-Features</b>	<b>0,79</b>	<b>0,80</b>	<b>0,86</b>	<b>3</b>

(a) Tabelle der Klassifikationsergebnisse



(b) ROC-Kurven der 15 Modelle

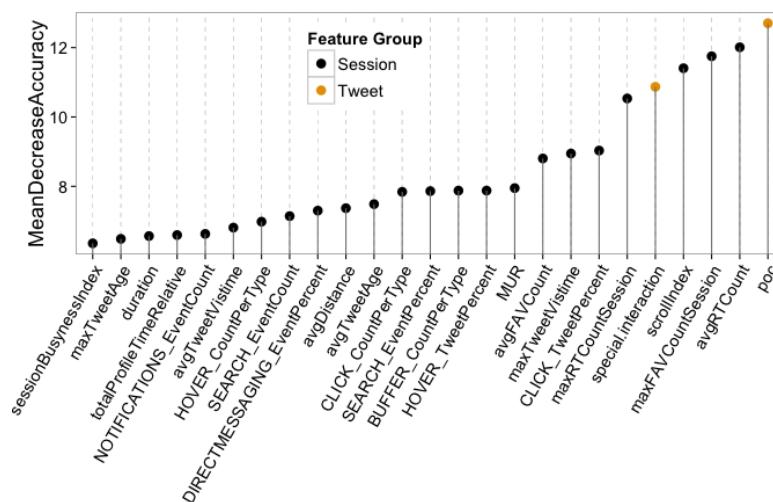
**Abbildung 6.30.:** Tabellarische und visuelle Darstellung (ROC-Kurven) der Klassifikationsergebnisse der 15 unterschiedlichen Modelle

plexität zu besitzen. Im Folgenden werden beide Perspektiven auf die Wahl der Features näher beleuchtet.

In der Praxis ist es notwendig, Klassifikatoren bzw. Modelle einsetzen zu können, die bei ihrer Vorhersage nicht auf Merkmale zurückgreifen, die auf der Erfassung und Speicherung des Langzeitnutzerverhaltens basieren, da dies zum einen infrastrukturell sehr aufwändig ist und zum anderen Vorhersagemöglichkeiten für alle Nutzer identisch möglich sein sollten und nicht nur für Nutzer, die bereits seit längerer Zeit aktiv sind. Lässt man die beiden Featuregruppen für welche dieser Sachverhalt zutrifft außer Acht — Nutzer-Merkmale und Merkmale der Nutzer-Sender-Interaktion — so lässt sich immer noch ein sehr gutes Klassifikationsergebnis erzielen. Dieses Modell nutzt Merkmale aus den Kategorien Tweet-Merkmale und Session-Merkmale und ist in Hinblick auf das F1-Maß das zweitbeste Modell (Tabelle 6.30a Zeile 2:  $F1 - Ma = 0,79$ ,  $Accuracy = 0,78$ ,  $AUC = 0,87$ ).

Insgesamt fließen in dieses Modell 97 unterschiedliche Variablen zur Klassifikation ein. Anders als bei einfachen Entscheidungsbäumen, bei denen der Baum visuell dargestellt werden kann und aus einzelnen Verzweigungen Regeln für heuristische Algorithmen abgeleitet werden können, ist die Interpretation von Random-Forest-Modellen nicht so geradlinig. Für ein besseres Verständnis, welche Merkmale zur Vorhersageleistung bzw. Klassentrennung beitragen, kann die sog. *variable importance* der Merkmale betrachtet werden. Hierzu kann die *mean decrease accuracy* als Maß herangezogen werden, welche zeigt, um wie viel die Genauigkeit der Vorhersageleistung abnehmen würde, wenn man die jeweilige Variable nicht in den Klassifikationsprozess mit einbezieht. Zur Berechnung der Accuracy wird der OOB-Fehlerwert (Out-of-bag-Fehler) genutzt. Durch das Bootstrapping-Verfahren werden beim Training eines jeden Baumes etwa  $\frac{2}{3}$  der Beobachtungen benutzt. Das verbleibende Drittel ist das sog. Out-of-bag-Sample der Beobachtungen, das als Testdatensatz interpretiert werden kann, anhand dessen die Fehlerrate des jeweiligen Baumes berechnet wird (James et al., 2013, S.317). Der OOB-Fehler ist äquivalent zu einer Leave-one-out-Kreuzvalidierung.

Abbildung 6.31 zeigt die 25 für die Klassifikationsleistung wichtigsten Merkmale. Zunächst ist zu erkennen, dass nur wenige Variablen sehr stark zur Klassentrennung beitragen — Top 6 — und die meisten Features nur einen sehr geringen Beitrag liefern. Darüber hinaus ist eine deutliche Dominanz von Session-Merkmalen zu erkennen, jedoch gehören auch zwei Tweet-Merkmale (POC



**Abbildung 6.31.:** Die Top 25 Merkmale basierend auf dem Maß *mean decrease accuracy* des Modells, welches auf Tweet- und Sessionmerkmale basiert. Die Merkmale sind nach Gruppenzugehörigkeit unterschiedlich farblich hervorgehoben.

und SPECIAL.INTERACTION) zur Gruppe der wichtigsten Features und sind damit gute Prädiktoren für ein zukünftiges Wiederfinden. Dass es sich dabei um die beiden Variablen POC (Tweet Teil einer Konversation) und SPECIAL.INTERACTION (Tweet favorisiert oder weitergeleitet) handelt, ist nicht überraschend, da beide Faktoren bereits vorher als relevante Charakteristika von wiedergefunden Tweets diskutiert wurden.

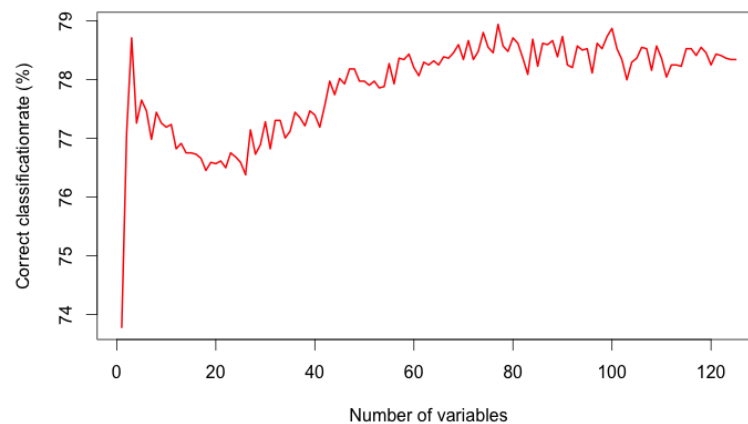
Eine weitere Herangehensweise besteht darin, den Merkmalsraum so niedrig wie möglich zu halten und damit auch die Rechenkomplexität für die Vorhersage zu beschränken. Für die Auswahl der wichtigsten Merkmale kann die *rfcv*-Funktion des R-Pakets *randomForest* genutzt werden, welche den Einfluss der einzelnen Merkmale für die Klassentrennung in einem kreuzvalidierten Verfahren bestimmt, wobei nach jedem Validierungsschritt die Accuracy der Klassifikation bestimmt wird und die Variable, welche den geringsten Einfluss auf die Klassentrennung besitzt, entfernt wird (Liaw & Wiener, 2002). Ausgehend von der Gesamtheit aller möglichen Merkmale (125) und damit auch den Features, die auf der Beobachtung des Langzeitinteraktionsverhaltens basieren, wird bestimmt, wie viele und welche Variablen für ein optimales Vorhersageergebnis benötigt werden.

Abbildung 6.32 zeigt den Prozentsatz der korrekt klassifizierten Instanzen (Genauigkeit) im Verhältnis zur Anzahl der für die Klassifikation genutzten Variablen. Es wird deutlich, dass mit drei Variablen ein ähnlich gutes Ergebnis bei der Klassifikation erreicht werden kann, wie dies mit 80 Variablen oder mehr der Fall ist. Der Umstand, dass drei Variablen ein besseres Klassifikationsergebnis erreichen als mehrere Variablen dies zu tun vermögen, begründet Breiman mit der Tatsache, dass sich Fehler vervielfachen können, wenn eine bestimmte Variable durch andere Variablen gestört wird bzw. durch diese das Rauschen erhöht wird (Breiman, 2001, S.24).

Diese drei Features sind:

- Anzahl der Sessions pro Nutzer (SESSIONCOUNT)
- Anzahl der Sessions pro Nutzer pro Tag (SESSIONSPERUSERPERDAY)





**Abbildung 6.32.:** Klassifikationsleistung (Accuracy) im Verhältnis zur Anzahl genutzter Variablen. Mit drei Top-Features kann ein identisches Ergebnis erreicht werden wie mit 80 Merkmalen oder mehr.

- Anzahl geklickter Tweets von Sender X (TWEETSClickedPerSender)

Alle drei Merkmale basieren auf der Beobachtung des Langzeitnutzerverhaltens und machen deutlich, dass häufige Twitter-Nutzung dazu führt, dass die Plattform auch als Informationsspeicher verwendet wird.

#### 6.4.5.4. Fazit zum Vorhersage-Experiment

Die letzten Abschnitte haben gezeigt, dass Wiederfindensziele mit einer angemessenen Genauigkeit vorhergesagt werden können. Die in Abschnitt 6.4.5.2 gebildeten und vorgestellten Features charakterisieren das Problem des Wiederfindens angemessen, da sämtliche Kombinationen aus Feature-Gruppen vergleichbare Ergebnisse liefern. Bereits das Baseline-Modell, also die Information, ob ein Tweet geklickt wurde und dessen Eigenschaften (Tweet-Features), sind ausreichend, um mit einer Genauigkeit von 72% vorherzusagen, ob ein Tweet wiedergefunden wird. Um eine höhere Vorhersagegenauigkeit zu erreichen, können zusätzliche Feature-Gruppen in den Klassifikationsprozess mit aufgenommen werden. Das beste Ergebnis lässt sich mit einem Modell aus 97 Merkmalen erreichen, welches die Gruppen Tweet-Features, Features, welche die Nutzer-Sender-Beziehung charakterisieren sowie die Session-Features kombiniert. Session-Features sind dabei besonders gute Prädiktoren. Jedoch auch die beiden Tweet-Features POC und SPECIAL.INTERACTION zeigen, dass Tweets, die Teil einer Konversation sind oder mit einem Relevanzkriterium (Favorisieren oder Weiterleiten) versehen werden, gute Prädiktoren sind. Mit Hilfe der `rfcv`-Funktion konnte ermittelt werden, dass bereits eine Kombination aus drei Features ausreicht, um eine sehr hohe Vorhersagegenauigkeit zu erreichen. Die Variable TWEETSClickedPerSender ist eine dieser drei Features. Sie modelliert die Empfänger-Sender-Beziehung, indem sie zeigt, wie viele Tweets eines Senders durch den Nutzer angeklickt werden. Wie in anderen Kapiteln unterstreicht dies die besondere Rolle, die dem Verhältnis zwischen Empfänger und Sender bei Praktiken des persönlichen Informationsmanagements auf Twitter zukommt. Das Vorhersageexperiment ist sicherlich noch nicht vollständig ausgeschöpft. Zum einen können andere Klassifikationsverfahren getestet

werden mit dem Ziel, die Vorhersagegenauigkeit zu optimieren. Besonders eine Feineinstellung der Parameter einer SVM könnten hier noch bessere Ergebnisse liefern. Zum anderen könnte eine Faktorenanalyse (Hauptkomponentenanalyse) genutzt werden, um den Feature-Raum zu verkleinern und Zusammenhänge zwischen Features noch deutlicher werden zu lassen.

## 6.5. Fazit zur Twitter-Logstudie

Die Twitter-Logstudie liefert sehr viele aufschlussreiche Erkenntnisse über das Nutzerverhalten beim Wiederfinden von Tweets. Viele der Verhaltensweisen, die sich in der qualitativen Studie bereits angedeutet haben, konnten quantitativ bestätigt werden. Nimmt man schließlich Bezug auf die eingangs formulierten Leitfragen, lassen sich folgende Ergebnisse festhalten:

Zum ersten Mal können auf der Basis einer wissenschaftlichen Studie Aussagen darüber getroffen werden, wie bzw. wodurch allgemeine Twitter-Nutzung charakterisiert ist. Twitter-Sitzungen sind im Durchschnitt etwa 4 Minuten lang. Die meisten Events entstehen dabei auf der eigenen *Timeline* während des Konsums der neuesten Tweets. Am zweithäufigsten werden die Profilseiten von anderen Nutzern besucht. Bei zunehmender Sitzungsdauer nimmt auch die Bedeutung weiterer Twitter-Funktionen, wie der Suche oder den Direktnachrichten, zu. Tatsächlich lässt sich auch in Bezug auf Twitter feststellen, dass nur wenige Interaktionen der Nutzer mit dem Social-Media-Service nach außen hin *sichtbare* Interaktionen — und damit mit Hilfe der API auslesbar — sind. Nur etwa 4% der betrachteten Tweets erhalten eine spezielle Interaktion. Dies bedeutet, dass auch auf Twitter ein Großteil der Interaktion mit dem Service *latentes* Verhalten, wie einfache Klicks auf Tweets oder der Besuch von Unterseiten, ist. In dieser Hinsicht bestätigen sich die Erkenntnisse von Schneider und Kollegen (Schneider et al., 2009). Bereits bei der allgemeinen Twitter-Nutzung lassen sich Indizien feststellen, die es erlauben, Twitter aus einer anderen Perspektive als der eines Echtzeitinformationsnetzwerkes zu betrachten. Die Daten zeigen, dass ein Tweet im Durchschnitt 28,1 Std. alt ist bis er auf der *Timeline* eines Nutzers erscheint, dass sehr viele ältere Tweets betrachtet werden und dass Nutzer dazu bereit sind, sehr viel Scrolling-Aufwand in Kauf zu nehmen, um zu bestimmten Tweets zu gelangen. All diese Faktoren deuten darauf hin, dass das Wiederfinden von Tweets ein gängiges und wiederkehrendes Informationsbedürfnis für Nutzer sein könnte.

Wie häufig tritt Wiederfinden tatsächlich auf? Diese Frage wurde in Abschnitt 6.4.4.2 diskutiert. Es wurde gezeigt, dass, wenn kein explizites Wissen in Form eines Goldstandards darüber vorliegt, ob Wiederfinden stattfindet oder nicht, auch eine Vielzahl an Platzhaltermaße vorhanden sind, die Wiederfindensverhalten approximiert. Legt man einen erneuten Klick in einer neuen Session als Platzhaltermaß fest, können 276 Re-finding-Aktionen bei 219 Tweets in 176 Sessions identifiziert werden. Da Tweets selten geklickt werden (nur 5% der betrachteten Tweets), ist dies ein sehr konservativer Proxy, der das tatsächliche Ausmaß an Wiederfindensverhalten vermutlich unterschätzt. Es konnte auch festgestellt werden, dass 55 Tweets häufiger als einmal wiedergefunden werden. Das Alter eines wiedergefundenen Tweets wurde im Rahmen der Leitfragen als der sog. Lebenszyklus eines Tweets bezeichnet und bezieht sich auf die zeitliche Dauer, die ein Tweet für einen Nutzer von Bedeutung/Relevanz sein kann und deswegen wiedergefunden wird. Im Durchschnitt werden die 219 Tweets nach 217 Std. (*Median* = 18,8 Std.) wiedergefunden. Der längste beobachtete Lebenszyklus eines Tweets liegt bei 52 Tagen, was bei einer Studiendauer von etwa fünf Monaten etwa die Hälfte der Studienzeit umfasst. Zudem konnte festgestellt werden, dass der Lebenszyklus stark davon abhängt, wie Nutzer im ersten Augenblick mit einem Tweet interagieren, da für Tweets, die nicht favorisiert oder weitergeleitet werden, die Wahrscheinlichkeit höher ist, erst zu einem sehr viel späteren Zeitpunkt wiedergefunden zu werden, wodurch der Lebenszyklus verlängert wird.

Eine der Hauptleitfragen aus Abschnitt 6.1 ist die Frage nach dem Verhalten beim Wiederfinden. Bereits auf Sessionebene lassen sich Unterschiede im Verhalten feststellen, da bei 28 von 58 Variablen, welche Sessionverhalten charakterisieren, deutliche Unterschiede in Re-finding-Sitzungen festgestellt werden können (vgl. Abschnitt 6.4.4.4). Sie sind zeitlich intensiver, besitzen mehr Events und mehr Klick-Interaktionen auf Tweets. Zudem befinden sich diese Tweets viel weiter unten in der *Timeline* und sind auch deutlich älter. Insgesamt wird deutlich, dass Wiederfindenssituationen durch anderes Verhalten charakterisiert sind. Eine weitere Leitfrage thematisierte die Strategien, die Nutzer verwenden, um zu Tweets zurückzukehren und welche Rolle vorheriges Aufbewahren für das Wiederfinden spielt. In Bezug auf die verwendeten Wiederfindensstrategien lassen sich drei wesentliche Feststellungen treffen. Wiederfinden findet nicht durch Eingabe einer Suchanfrage statt. Schlüsselwortsuche spielt beim Wiederfinden von Tweets eine untergeordnete Rolle. Wiederfindensstrategien, die ein vorheriges Aufbewahren von Tweets verlangen — Wiederfinden über die Favoritenliste oder die eigenen gesendeten Tweets via RT — werden zwar sowohl in kurzen und langen Wiederfindensperioden gleich häufig genutzt, insgesamt betrachtet werden sie aber ebenfalls kaum verwendet. Die häufigsten Wiederfindensaktionen finden auf Profilseiten von anderen Nutzern statt. Hierbei lassen sich in Anlehnung an bestehende PIM-Literatur zwei Verhaltensweisen identifizieren: Orientierungsverhalten, das auf Suche aufbaut und linkbasiertes Orientierungsverhalten.

Abschließend wird in einem Experiment untersucht, ob es möglich ist, potentielle Wiederfindensziele vorherzusagen. Hierfür werden mittels Feature-Engineering 125 Features erstellt und in einem Klassifikationsszenario verwendet. Bereits einfache Tweet-Features reichen dazu aus, um mit einer Genauigkeit von 72% vorherzusagen, ob ein Tweet wiedergefunden wird oder nicht. Reduziert man den Feature-Raum wird deutlich, dass häufige Twitter-Nutzung und das Verhältnis zum Tweet-Sender eine wesentliche Rolle im Vorhersageprozess darstellen. Nimmt man Bezug auf die Frage, ob Tweets bestimmte Charakteristika besitzen, die ein Wiederfinden wahrscheinlicher machen, lässt sich lediglich die Tatsache, ob ein Tweet Teil einer Konversation ist (also mit einer *@Mention* beginnt), als Charakteristikum identifizieren, das als Indiz für ein späteres Re-finding gelten kann. In einem Kodierungsprozess, welcher wiedergefundene Tweets in Kategorien einteilt, kann keine Tweet-Kategorie ausgemacht werden, die übermäßig häufiger wiedergefunden wird.

Trotz vieler wesentlicher Erkenntnisse muss auch Kritik geübt werden bzw. können Verbesserungsmöglichkeiten aufgezeigt werden. Ein wesentlicher Kritikpunkt liegt in der Tatsache begründet, dass — bis auf eine Ausnahme — nur *erfolgreiches* Wiederfinden analysiert wird und schwierige oder frustrierende Wiederfindenssituationen nicht behandelt werden. Vielmehr bestand durch die Art der Datenerhebung keine Möglichkeit, erfolgloses Wiederfinden konkret zu erfassen. Somit basieren die Ausführungen zum Lebenszyklus eines Tweets oder dem Wiederfindensverhalten im Speziellen lediglich auf erfolgreichen Situationen, in denen das Wiederfindensziel gefunden und das Informationsbedürfnis befriedigt wurde. Lediglich in Abschnitt 6.4.4.8 wurde versucht, sich diesem Problem des erfolglosen Wiederfindens mit Hilfe von *Proxies* und dahinterstehenden Annahmen zu nähern und somit erfolglose Wiederfindenssituationen zu charakterisieren. Hier besteht Raum zur Verbesserung, welche in diesem Fall im Bereich der Datenerhebung und somit in den Interaktionsmöglichkeiten mit der Chrome-Erweiterung liegt. Eine Möglichkeit besteht in der Ergänzung der Chrome-Erweiterung um eine sogenannte *browser action*, wie sie in Abschnitt 6.3.1 beschrieben wurde, mit deren Hilfe der Nutzer Wiederfindensaktionen per Klick manuell angeben kann. So könnte der Browser mit einem Icon/Button versehen werden, welcher stets dann betätigt wird, wenn eine Re-finding-Aktion beginnt und erneut betätigt wird, wenn die Suche nach dem Wiederfindensziel beendet ist. Auf diese Weise könnten auch erfolglose Wiederfindensaktionen explizit registriert werden. Eine vergleichbare Herangehensweise findet sich bei Elsweiler und Kollegen, welche in einer Logstudie zum Wiederfindensverhalten bei E-Mails einen E-Mail-Client um diese Funktionalität erweitern (Elsweiler, Harvey & Hacker, 2011). Jedoch stellt auch diese

Methode keine vollkommene Lösung dar, da Nutzer sich zum einen daran erinnern müssen, die Buttons zu klicken und zum anderen bei bestimmten Umständen z. B. bei einfachen, schnellen Wiederfindensaktionen Nutzer davon absehen könnten, dies zu tun.

In den folgenden beiden Kapiteln werden die Erkenntnisse der beiden Studien vor dem Hintergrund der bestehenden Literatur diskutiert sowie mögliche Lösungen für die Unterstützung des Nutzers beim Wiederfinden aufgezeigt.

# 7

## **Diskussion: Erkenntnisse aus Literatur und Studien vereint**

Kapitel 7 stellt die Phase der Meta-Inferenz des in Abschnitt 4.3 vorgestellten Sequential-Mixed-Methods-Design dar. Ziel der Meta-Inferenz ist es, die Erkenntnisse der beiden Studien vor dem Hintergrund des existierenden Forschungsstandes und den formulierten Leitfragen zusammenzufassen und zu diskutieren. Aufgrund der verschiedenen Datenquellen erfolgt somit sowohl eine subjektive als auch objektive Charakterisierung des Wiederfindensverhaltens auf Twitter auf der Basis von qualitativen und quantitativen Daten und deren Interpretation. Dabei wird auch auf die Diskrepanz zwischen dem, was Nutzer glauben zu tun bzw. sagen, dass sie es tun und ihrem tatsächlichen Verhalten eingegangen. Im Folgenden wird anhand der beiden PIM-Praktiken Aufbewahren und Wiederfinden eine Zusammenfassung der Forschungsfragen aus den einzelnen Unterkapiteln vorgenommen, wobei diese jeweils vor dem Hintergrund der Literaturlage, den Ergebnissen der Umfragestudie sowie den Ergebnissen der Logstudie beantwortet werden.

### **7.1. Fragestellungen zur PIM-Aktivität Aufbewahren**

Aufbewahren und Wiederfinden sind komplementäre Handlungen, weshalb das Aufbewahren einen wesentlichen Bestandteil des Wiederfindensverhaltens darstellt. Abschnitt 2.2.2 hat eine detaillierte Analyse dieser PIM-Praktik vorgenommen und festgestellt, dass dieses Verhalten sehr individuell sein kann und zahlreiche Faktoren — angefangen vom Medientyp bis hin zum zukünftigen Nutzungskontext — Einfluss darauf haben, wie Aufbewahrensentscheidungen getroffen werden und welche Handlungen aus diesen Entscheidungen resultieren. Aufbewahren ist kognitiv fordernd, da es vom Nutzer einerseits einen Abgleich zwischen aktuell vorliegender Information und zukünftigem Informationsbedürfnis — also den vielfach angesprochenen Blick in die Zukunft — verlangt und der Nutzer stets Faktoren wie die Kosten des Aufbewahrens, sowie die Kosten, eine Information in Zukunft nicht vorliegen zu haben, abwägen muss.

### 7.1.1. Aufbewahrensbedürfnis und Strategien des Aufbewahrens

Vor Studienbeginn existierte kaum wissenschaftliche Evidenz darüber, ob Nutzer das Bedürfnis haben, Tweets aufzubewahren. Die Twitter-Umfrage hat jedoch offenbart, dass über dreiviertel der Probanden (77,3%) schon einmal das Bedürfnis hatten, einen Tweet zu speichern. Die Korrelationsmatrix aus Abbildung 5.23 zeigt zudem, dass Nutzer, die häufig aufbewahren auch dazu tendieren, häufiger wiederzufinden ( $\rho = 0,36$ ), häufiger eigene Tweets wiederzufinden ( $\rho = 0,36$ ) und stärker der Aussage zustimmen, bereits einmal bei einem Wiederfindensprozess frustriert worden zu sein ( $\rho = 0,34$ ). Es ist nur konsequent, dass die Variable FREQUENCY: PRESERVING TWEETS einen höchst signifikanten Prädiktor im Regressionsmodell bei der Vorhersage von Frustration beim Wiederfinden bildet. Vergleicht man die Zahl der Probanden, die angeben, Tweets aufzubewahren, mit der Zahl der Probanden, die angeben, Tweets wiederzufinden, so liegt diese Zahl etwas höher. Eine mögliche Erklärung hierfür wird in Abschnitt 7.1.2 gegeben.

Auch die Frage nach den verwendeten Aufbewahrensstrategien war folglich unbeantwortet. Eine Studie zum RT-Verhalten von boyd, Golder und Lotan offenbart die Aufbewahrens-Wiederfindens-Kombination aus dem Weiterleiten eines Tweets, um ihn später in den eigenen gesendeten Tweets einfacher lokalisieren zu können (boyd et al., 2010). Favorisieren, um Tweets aufzubewahren bzw. ein Lesezeichen zu setzen, wird in späteren Arbeiten, wie von Gorrell und Bontcheva, näher exploriert (Gorrell & Bontcheva, 2016). Zum Zeitpunkt der Konzeption der Twitter-Umfrage existierten nur anekdotische Hinweise auf Blogs, Foren und Webseiten — unter anderem auch der Twitter FAQ — die darauf hindeuteten, dass auch andere Twitter-Funktionen, wie die Favoriten-Funktion, zum Setzen von Lesezeichen genutzt werden (Stone, 2006). Tatsächlich stimmen fast 50% der Probanden der Aussage zu, die RT-Funktion zum Aufbewahren von Tweets zu nutzen. Ein beinahe ebenso großer Prozentsatz stimmt der Aussage zu, Tweets zu favorisieren (46%), um diese in der Favoritenliste zu speichern. Lediglich 29% geben an, externe Speicher zu verwenden. Gruppiert man die Probanden nach Teilnehmern, die sich zufrieden über die twitterinternen Funktionen zum Aufbewahren von Tweets äußern, so kann man erkennen, dass diese der Nutzung von Weiterleiten und Favorisieren zum Aufbewahren signifikant stärker zustimmen als Nutzer, die sich negativ über die Funktionen äußern (vgl. Abschnitt 5.7.2 und Abbildung 5.17).

Aufbewahrensverhalten ist in den Logdaten weitaus schwieriger zu identifizieren und analysieren. Aus den Logs können keine Nutzerabsichten abgeleitet werden, so dass nur Platzhaltermaße als Interpretationsansätze bei der Beschreibung des Verhaltens genutzt werden können. Dies wird vornehmlich in Abschnitt 7.2.1 erfolgen.

Dass die Funktion des Favorisierens eine wichtige Bedeutung für das PIM-Verhalten auf Twitter besitzt, wird in mehreren Abschnitten der Arbeit deutlich. Besonders evident wird es bei der Darstellung der Gründe für das Favorisieren — insgesamt wurden 25 Beweggründe ermittelt — da die Nutzung der Funktion als Bookmarking-Werkzeug, die zweithäufigste Begründung für die Verwendung der Funktion darstellt (vgl. Abschnitt 5.6.3.2). Die Tatsache, dass Kode [B1] *Bookmarking* aber noch sechs Sub-Kodes besitzt, zeigt, wie differenziert das Aufbewahren von Nutzern betrachtet wird. Dass Nutzer Tweets favorisieren, um sie später erneut aufzusuchen, Handlungen (*future use*) vorzunehmen und dabei oft schon ein zukünftiges Informationsbedürfnis antizipieren, passt in die Interpretation des Aufbewahrens als Komplementärhandlung für ein späteres Wiederfinden. Die Sub-Kodes des Kodes [B1.2] *Memory* zeigen jedoch, dass Nutzer Tweets auch aus sentimentalen Gründen bzw. der Erinnerung wegen aufbewahren. Wie diese sentimental Motive und die Wahrnehmung von Twitter als Archiv zusammenhängen wird im Folgenden beschrieben.

### 7.1.2. Twitter als Archiv

Eine Leitfrage in Abschnitt 5.1 war die Frage, ob Twitter Archivfunktion besitzt bzw. von seinen Nutzern als Archiv wahrgenommen wird. Diese Annahme beruht auf Erkenntnissen von Arbeiten zu anderen Social-Media-Plattformen, die in Abschnitt 3.2.2.1 vorgestellt wurden. Die dort diskutierten Arbeiten von Zhao et al., Marshall und Shipman, Zarro, Hall und Forte oder Linder, Snodgrass und Kerne verdeutlichen, dass verschiedene Social-Media-Applikationen wie Facebook oder Pinterest als Archive wahrgenommen werden, die dazu genutzt werden, um persönliche Kollektionen von Nachrichten zu erstellen und in Erinnerungen zu schwelgen (Zhao et al., 2013; Marshall & Shipman, 2014; Zarro & Hall, 2012; Linder et al., 2014). Für die Twitter-Nutzung steht sicherlich der Echtzeitcharakter im Vordergrund. Wie bereits gesehen, lassen sich auch Spuren erkennen, die darauf hindeuten, dass auch Twitter eine gewisse Archivfunktion besitzt.

Weiter oben wurde angesprochen, dass mehr Nutzer das Bedürfnis verspüren, Tweets aufzubewahren, als diese wiederzufinden (77,3% vs. 72,4%). Dies kann unter anderem dadurch begründet werden, dass Aufbewahren nicht immer rational motiviert sein muss, sondern es viele weitere Motivationsgründe gibt — wie auch die Gründe für das Favorisieren in Kode[B1.2] zeigen — die den Nutzer zum Aufbewahren motivieren können. Motive wie „So i can always remember the tweet“ [R13] oder „Quotes or links that i want to remember“ [R133] geben Hinweise darauf, dass Nutzer eine Sammlung interessanter Tweets anlegen, deren Zweck nicht im einmaligen Wiederfinden besteht, sondern die Twitter bzw. der Favoritenliste eine Art Archivfunktion zuschreiben.

Ferner geben auch die Wiederfindensgründe Aufschluss über den Archivcharakter Twitters. Dort ist zu erkennen, dass Nutzer aus sentimental oder emotionalen Beweggründen heraus Tweets wiederfinden wollen und hierbei oftmals mehr als nur einen Tweet betrachten. Besonders evident wird dies in der Verhaltensweise, die in dieser Arbeit — inspiriert durch Nutzerkommentare — als *Timeline Strolling*, also als Spaziergang durch die *Timeline*, bezeichnet wird: „Other times it’s just strolling back through my time line to find things I’d retweeted just to enjoy them again (usually not a specific tweet in mind)“ [R54]. Eigene gesendete Tweets werden aus Nostalgie heraus erneut betrachtet, wobei Nutzerkommentare auch darauf hindeuten, dass bei dieser Betrachtung eine gewisse Selbstreflexion stattfindet und Nutzer via Evaluation ihres Tweet-Verhaltens zu unterschiedlichen Lebenszeitpunkten zurückkehren und über ihre eigene Entwicklung nachdenken.

Auch rein objektiv lässt sich feststellen, dass Twitter Archivcharakter besitzt und nicht nur Echtzeitinformationen konsumiert werden. Evident wird dies, wenn man das Alter der Tweets betrachtet, die während der Logstudie auf den Bildschirmen der Probanden erscheinen. Das relativ hohe Durchschnittsalter (*Mittelwert* = 28,1 Std., *Median* = 1,4 Std., *Max* = 80667 Std.) der Tweets mit denen Nutzer interagieren, gibt nicht nur Hinweise darauf, dass Nutzer das Bedürfnis haben, ältere Tweets zu lesen und diese möglicherweise hierfür wiederfinden, sondern zeigt zudem auch, dass auf Twitter nicht nur Echtzeitinformation konsumiert wird, sondern Informationen auch für längere Zeit gespeichert sind und wieder abgerufen werden. Dies zeigt sich auch dadurch, dass Nutzer während der Studie sogar zum aller ersten gesendeten Tweet von Jack Dorsey zurückkehren. Die Erkenntnis von Wu et al., dass bestimmte Inhalte auf Twitter eine beinahe unendlich lange Lebensphase aufweisen und ständig auf der Plattform zu finden sind, deutet darauf hin, dass Twitter ein Archiv sein kann (Wu et al., 2011). Aber nicht alle Tweets sind für immer zugänglich. In einer aktuellen Studie von Mondal et al. zum Longitudinal-Privacy-Verhalten stellen diese fest, dass in ihrem Datensatz, bei Tweets die sechs Jahre oder älter waren, etwa ein Drittel der Nachrichten nicht mehr zugänglich war (Mondal et al., 2016).

Zum Abschluss der Diskussion, ob Twitter als Archiv wahrgenommen wird, wird eine Aussage aus den Wiederfindensgründen erneut zitiert: „Because they contain some relevant information that I don’t remember; it is very much like knowing that an information is written down in a specific boo-

k/chapter, and you can look it up whenever you need to“ [R202]. Vermutlich ist die Wahrnehmung Twitters als immanentes, allspeicherndes und zugänglichmachendes Medium, wie es im letzten Kommentar beschrieben wird, nicht als weitverbreitet anzusehen, jedoch scheint diese Auffassung durchaus vorhanden zu sein.

## 7.2. Fragestellungen zur PIM-Aktivität Wiederfinden

Eine wesentliche Frage, die in beiden Studien formuliert wurde, ist die Frage nach der Frequenz des Wiederfindens. Wie häufig geben Nutzer an, das Bedürfnis zu haben, einen Tweet wiederfinden zu müssen und wie steht diese Aussage im Verhältnis zum tatsächlichen Wiederfindensverhalten, das objektiv in den Logdaten beobachtet werden kann?

In der Umfrage geben 72,4% aller Probanden an, dass sie schon einmal einen Tweet wiederfinden wollten bzw. erneut aufgesucht haben. Entsprechend kann etwa bei einem Viertel davon ausgegangen werden, dass sie dieses Bedürfnis noch nie hatten. Dass die Zahl der Personen, die wiederfinden geringer ist als die Zahl der Probanden, die Tweets aufbewahren, wurde bereits angesprochen. Bei der Betrachtung der Gruppe der Personen, die häufiger wiederfinden, stellt man fest, dass 8,4% der Personen angeben, mindestens wöchentlich Tweets erneut auffinden zu wollen. Betrachtet man das Verhältnis der Wiederfindensfrequenz mit anderen Variablen wird deutlich, dass Nutzer, die häufig wiederfinden, dies auch häufig bei eigenen gesendeten Tweets tun ( $\rho = 0,36$ ) — weiter unten werden mögliche Gründe vorgestellt, die das Wiederfinden von eigenen Tweets motivieren können — und dass Nutzer, die häufig wiederfinden, auch häufiger bereits einmal ein frustrierendes Wiederfindenserlebnis erfahren haben ( $\rho = 0,36$ ). Diese hohen Korrelationen spiegeln sich auch im logistischen Regressionsmodell wider, da die Häufigkeit des Wiederfindens — sowohl der eigenen gesendeten Tweets als auch der Tweets aus der Timeline — dort höchst signifikante Prädiktoren für ein frustrierendes Wiederfindenserlebnis darstellen (vgl. Abschnitt 5.7.8).

Der Frage nach der Auftretenshäufigkeit von Wiederfindensaktionen in den Clickstream-Daten ist die Frage vorgelagert, wie Wiederfinden definiert wird bzw. wie Wiederfindensaktionen in Logdaten identifiziert werden können. Abschnitt 6.4.4.2 nimmt eine Analyse von Verfahrensweisen vor, die in der Arbeit genutzt werden hätten können, um Wiederfinden zu identifizieren. Die Verfahrensweisen sind dabei auf einem Kontinuum verortbar, das von liberal und Recall-orientiert — möglichst alle Wiederfindensaktionen zu erhalten — bis konservativ und Precision-orientiert — höhere Sicherheit, dass die Instanzen, die man diskutiert auch tatsächlich Wiederfindensaktionen sind — reicht. Die Verfahren werden anhand wesentlicher Charakteristika wie *Anzahl der Sitzungen in denen Wiederfinden stattfindet*, *Anzahl der Personen, die Wiederfinden* und *Anzahl der wiedergefundenen Informationsobjekte* verglichen. Für die weiteren Analysen wird ein erneuter Klick auf einen Tweet in einer neuen Session als Platzhaltermaß für eine Re-finding-Aktion gewertet, wie dies auch bei Wiederfindensstudien in anderen Kontexten der Fall ist (Teevan et al., 2006; Harvey & Elweiler, 2012). Dies ist ein sehr konservativer Proxy, der das Vorkommen von Wiederfinden in den Logs unterschätzt. Er garantiert jedoch eine hohe Genauigkeit/Sicherheit. Wendet man den Proxy Across-Session-Re-Klick auf die Logdaten an ergeben sich:

- 276 Wiederfindensaktionen (0,2% aller Events)
- 219 Tweets, die wiedergefunden werden (0,2% aller gesichteten Tweets)
- 176 Wiederfindenssitzungen (3,9% aller Sitzungen)
- 17 Nutzer, die wiederfinden (38,6% aller Nutzer)

Der Anteil der Probanden, die in der Twitter-Umfrage angeben, Tweets wiederzufinden, liegt deutlich unter dem Anteil der Nutzer, die in der Logstudie tatsächlich dieses Verhalten an den Tag legen



(72,4% vs. 38,6%). Dies kann an der Wahl des konservativen Proxies liegen, bzw. an der Tatsache, dass grundsätzlich mit 5,7% (vgl. Tabelle 6.3) nur wenige Tweets überhaupt geklickt werden. Ferner ist ein Grund sicherlich die Tatsache, dass Nutzer Twitter von mehreren Endgeräten aus nutzen und Cross-device-Interaktionen nicht berücksichtigt werden konnten. Es ist fraglich, inwieweit die Ergebnisse von Studien, die das Wiederfindensverhalten in anderen PIM-Kontexten und mit anderen Informationsobjekten untersuchen, überhaupt als Vergleichsparameter herangezogen werden können. Betrachtet man beispielsweise die Ergebnisse von Obendorf et al., die feststellen, dass 41% aller Website-Aufruf Wiederaufrufe sind (Obendorf et al., 2007), von Teevan et al., die 39% aller Suchanfragen als Re-finding-Suchanfragen identifizieren (Teevan et al., 2006), oder von Elswiler, Harvey und Hacker, die zeigen, dass 55% aller E-Mails erneut geöffnet werden, so liegen diese sehr deutlich über der Anzahl der Wiederfindensaktionen in dieser Arbeit. Dies ist vermutlich nicht weiter verwunderlich, da der Fokus Twitters klar auf Echtzeitinformationen liegt. Jedoch ist es für ein Echtzeitinformationsnetzwerk erstaunlich, dass dieses als Archiv wahrgenommen wird und dass mit einem hohem Ausmaß Nutzer das Bedürfnis haben, zu bereits gesehenen Tweets zurückzukehren. Im Folgenden werden die Wiederfindensgründe der Twitter-Umfrage mit Wiederfindensgründen aus der Logstudie kontrastiert.

### Wiederfindensgründe im Vergleich

Die Freitextfragen der Twitter-Umfrage haben eine große Bandbreite an Wiederfindensgründen offenbart (vgl. Abschnitt 5.8.1). Sie reichen von Wiederfinden einfacher Zitate, Links oder Bilder, über Wiederfinden, um bestimmte Handlungen mit dem Tweet vorzunehmen, wie ihn mit anderen Personen zu teilen, bis hin zu sentimental Motiven, wie das Betrachten eigener gesendeter Tweets aus Nostalgie. Wie bereits in Abschnitt 2.1.3, bei der Darstellung der Charakteristika von Tweets beschrieben wurde, ist eine Aufteilung in Wiederfindensgründen aus einem Handlungsbedarf heraus (*actionable*) und aus einem reinen Informationsbedarf (*informative*) zu erkennen. Überraschend ist die Häufigkeit mit der Nutzer angeben, dass nicht nur einzelne Tweets das Wiederfindensbedürfnis stillen können, sondern, dass mehrere Tweets — z. B. alle Tweets einer Konversation oder alle Bilder, die ein bestimmter Nutzer gesendet hat — besucht werden. Dies erklärt auch einige Verhaltensweisen, die in der Logdatenstudie beobachtet werden können. So zeigt der Boxplot in Abbildung 6.21, welche die Zeitperiode zwischen ersten und erneutem Klick auf einem Tweet visualisiert, Häufungen von erneuten Klicks bei 505 Std. und bei etwa einem Monat, welche als Multi-item-Re-finding interpretiert werden können.

Vergleicht man die Wiederfindensgründe aus der qualitativen Auswertung mit den Wiederfindensgründen der Logstudie, die darauf basieren, dass der Eventtyp, der für den erneuten Klick verantwortlich ist, als Proxy für den Wiederfindensgrund gewertet wird (vgl. Abbildung 6.26), so zeigt sich eine hohe Übereinstimmung. Die Klick-Ereignisse der erneuten Klicks lassen sich gut als Platzhaltermaße verwenden, um zu erkennen, wie Wiederfinden motiviert ist. Etwa die Hälfte aller wiedergefundenen Tweets der Logstudie sind als *Media Re-finding* zu interpretieren, bei dem Nutzer das Ziel hatten, den Link oder das Bild/Video/Website erneut zu besuchen. In 11% der Fälle ist das Teilen auf der Plattform, also der RT, der Wiederfindensgrund. In 6,5% ist es das private Weiterleiten. Beides entspricht den *Social Motives* des Kodierschemas der Wiederfindensgründe. Bei knapp 10% der Tweets war der Wiederfindensgrund das Verfassen einer Antwort, was in der qualitativen Analyse dem Kode [1.1] *Administrative Motive: Tweet Composing* entspricht. Unterstellt man Nutzern, die einen Tweet beim erneuten Besuch favorisieren ein Aufbewahrensmotiv — was durchaus plausibel erscheint, da einige Nutzer das Aufbewahren von Tweets als Wiederfindensgrund angeben — so entspricht dies dem Kode [1.3] *Specific Use*. Es wird evident, dass die Wiederfindensgründe beider Studien sehr kongruent sind.

In Abschnitt 3.2.2 wurde ausführlich das Verhalten der *regret deletion* vorgestellt, welches in zahlreichen Studien untersucht wird und durchaus häufig auftritt. Interessanterweise enthielt kei-

ner der Kommentare, die für die Analyse der Wiederfindensgründe in Abschnitt 5.8 ausgewertet wurden, den Term *delete* oder das Löschen als Absicht. Löschungen in den Logdaten (Eventtyp: *TweetDelete*) wurden registriert aber nicht näher in den Studien untersucht, da der Fokus der Arbeit auf Tweets aus der *Timeline* liegt.

### 7.2.1. Relevanz von Tweets: Jetzt! Und in Zukunft?

In den Twitter-Arbeiten, die in Abschnitt 3.2.1 zum Thema Microblogsuche bzw. Auffinden von Social-Media-Nachrichten vorgestellt wurden, wurde ein RT oder ein Favorisieren häufig als Relevanzkriterium erachtet. Wird ein Tweet häufig weitergeleitet oder favorisiert, so hat er immer auch eine gewisse Relevanz für eine breitere Gruppe (global gesehen über alle Nutzer auf der Plattform hinweg), oder auch für individuelle Nutzer. Dort wurden auch Aussagen über den globalen Lebenszyklus von Tweets getroffen. Kwak et al. stellen fest, dass 10% aller RTs noch nach einem Monat stattfinden (Kwak et al., 2010). Wu et al. erkennen, dass bestimmte Themen eine beinahe unbegrenzte Lebensdauer besitzen, wobei diese weniger durch RTs als durch ein erneutes Entdecken bzw. Auftreten in neuen Tweets zu erklären ist (Wu et al., 2011). Vergleichbare Studien zu anderen Microblog-Plattformen wie Tumblr äußern sich konträr gegenüber der Lebensdauer von Inhalten auf deren Plattform. Zum einen wird argumentiert, sie sei länger als auf Twitter (J. Xu et al., 2014), zum anderen wird jedoch ausgesagt, sie sei deutlich kürzer (Chang et al., 2014). Wie steht es um die Lebensdauer von Tweets aus individueller Perspektive? Mit anderen Worten: Wie lange können Tweets für Nutzer von Relevanz sein? Abbildung 6.21 in Abschnitt 6.4.4.3 zeigt, dass der Lebenszyklus von Tweets relativ kurz ist. Im Median werden Tweets nach 18,8 Std. (1. *Quantil* = 0,5 Std., *Mittelwert* = 217 Std., *Max* = 1243,8 Std. 52 Tage) wiedergefunden, also etwa innerhalb eines Tages. Das Maximum von 52 Tagen zeigt jedoch, dass Tweets auch über mehrere Wochen hinweg noch von Relevanz für den Nutzer besitzen können. Etwa ein Viertel aller Tweets wurde innerhalb einer Stunde wiedergefunden, ein Wert der vergleichbar ist mit der Wiederfindensperiode von E-Mails bei Elweiler und Harvey, die für 36% aller E-Mails eine Stunde beträgt (Elweiler, Baillie & Ruthven, 2011). Kooti et al. stellen in einer aktuellen Logstudie mit 16 Milliarden E-Mails — interpretiert man die Antwort (*Reply*) auf eine E-Mail als Wiederfinden — fest, dass 90% aller E-Mails innerhalb einer Stunde beantwortet werden, wobei aber die Hälfte aller Antworten erst nach 47 Minuten gegeben werden (Kooti et al., 2015). All diese Verteilungen sind sehr ähnlich zur von Hodas und Lerman festgestellten Zeitperiode, die verstreicht, bis Nutzer einen Tweet weiterleiten (Hodas & Lerman, 2012). Es zeigt sich, dass trotz Twitters Echtzeitcharakters zeitliche Muster zu erkennen sind, die sehr ähnlich auch beim Verhalten mit E-Mail auftreten. Betrachtet man im Folgenden den Anteil der wiedergefundenen Tweets, die eine spezielle Interaktion erfahren haben — also ein Relevanzsignal — so zeigt sich, dass der Anteil der Tweets mit 61,6% in der kurzen Wiederfindensperiode ( $< \text{Median} = 18,8 \text{ Std.}$ ) deutlich höher ist als in der langen Wiederfindensperiode, in welcher der Anteil an favorisierten und/oder weitergeleiteten Tweets nur bei 10,4% liegt. Legt man wieder die Annahme zu Grunde, dass relevante Tweets aufbewahrt werden und aus diesem Grund eine spezielle Interaktion erfahren, so müsste der Anteil jedoch deutlich größer sein, da Nutzer die RT- bzw. FAV-Funktion als Aufbewahrensmechanismen nutzen. Das *odds ratio* eines signifikanten Chi-Quadrat-Tests zeigt jedoch, dass die Wahrscheinlichkeit, dass ein Tweet nach langer Periode erneut geklickt wird, für Tweets ohne spezielle Interaktion 14-fach höher liegt. Einen möglichen Erklärungsansatz für diese Beobachtungen bietet die mehrfach in Kapitel 2 angesprochene PAIN-Hypothese, die besagt, dass sich Nutzer in Bezug auf Information stets mit einer Aufbewahrensentscheidung konfrontiert sehen und hierfür ein Blick in die Zukunft verlangt wird bzw. zukünftige Informationsbedürfnisse vorhergesagt werden müssen (Bruce, 2005). Dass der Anteil an Tweets ohne spezieller Interaktion in der Gruppe der nach langer Periode wiedergefundenen Tweets so hoch ist, deutet darauf hin, dass Nutzer Probleme haben, ein *mapping*

zwischen Tweets und Informationsbedürfnissen herzustellen bzw. es ihnen nicht gelingt, zukünftige Informationsbedürfnisse vorherzusagen. Stattdessen kommt es in Bezug auf Tweets häufiger vor, dass deren Wert erst nach längerer Zeit erkannt wird, so dass der von Wen festgestellte Zustand des *post retrieval value*, also die nachträgliche Erkenntnis der Relevanz eines Informationsobjekts, auch bei Tweets beobachtet werden kann (Wen, 2003). Für den individuellen Lebenszyklus bedeutet dies schließlich: Favorisieren und Weiterleiten verkürzen den nutzerbezogenen, persönlichen Lebenszyklus.

Wenn Nutzer nicht oder nur eingeschränkt über die Fähigkeit verfügen, zukünftige Informationsbedürfnisse vorherzusagen, stellt sich die Frage, ob ein Algorithmus dies an deren Stelle tun kann und sie somit bei Wiederfinden unterstützen kann. Diese Frage wurde als eine Leitfrage der Twitter-Logstudie formuliert und in Abschnitt 6.4.5 in Form von Machine-Learning-Experimenten exploriert. Die Experimente sind als binäre Klassifikationsexperimente konzipiert, bei denen vorhergesagt werden soll, ob ein Tweet wiedergefunden wird oder nicht. Für die Vorhersage werden 125 Features genutzt, welche in Tabelle 6.7 aufgeführt sind. Diese können weiter in folgende vier Gruppen eingeteilt werden:

- **Tweet-Features:** Charakterisieren den Tweet, dessen Metadaten (z. B. Anzahl der Favorisierungen) und den Tweetinhalt durch Anzahl der Zeichen oder Wörter, Anzahl enthaltener Sonderzeichen (Hashtags, Mentions) usw.
- **Session-Features:** Diese Featuregruppe charakterisiert das Verhalten des Nutzers in der Session, in welcher der Tweet zum ersten Mal geklickt wurde (z. B. Dauer der Session und Anzahl an Events).
- **Features zur Nutzer-Sender-Beziehung:** Diese Features erfassen das Verhältnis des Rezipienten und des Tweet-Senders/Autors (z. B. Anzahl geklickter Tweets, die von einem Sender gepostet wurden).
- **Nutzer-Features:** Sie beschreiben das Nutzungsverhalten der Probanden (z. B. Anzahl der Sessions pro Tag).

Tabelle 6.30a in Abschnitt 6.4.5.3 zeigt, dass die Vorhersagegüte aller 15 möglichen Modelle durchaus zufriedenstellend ist. Sogar das Baseline-Modell, also dasjenige Modell, das nur Tweet-Features benutzt, die leicht über die API zu erhalten sind, reicht dazu aus, um mit einer Genauigkeit von 72% vorherzusagen, welcher Tweet wiedergefunden wird. Abbildung 6.31 gibt zusätzlich darüber Aufschluss, welche Charakteristika wiedergefundene Tweets besitzen. Dort sind zwei Tweet-Features aufgeführt, nämlich *SPECIAL.INTERACTION*, also ob mit einem Tweet speziell interagiert wurde und *POC*, also ob ein Tweet Teil einer Konversation war, also mit einer *@Mention* beginnt. Andere Charakteristika wie Tweetlänge, Art des Tweets oder Lesbarkeit spielen eine untergeordnete Rolle. Unter Verwendung eines Feature-Reduktions-Algorithmus (*rfcv*-Funktion) konnte zusätzlich festgestellt werden, dass mit nur drei Features ein ähnlich gutes Ergebnis erreicht werden kann, als wenn einem Algorithmus 80 oder mehr Features zur Verfügung stehen. Dies ist ein wünschenswerter Zustand, da somit für eine Vorhersage weniger Daten gespeichert und für die Berechnung verarbeitet werden müssen, was in der Praxis eine zeitnahe Vorhersage der Wiederfindensziele garantiert. Eines der drei Features beschreibt das Nutzer-Sender-Verhältnis, also die Anzahl der Tweets eines Senders, die ein Rezipient klickt. Es wird deutlich, dass Algorithmen besser in der Lage sind, Wiederfinden bzw. zukünftige Informationsbedürfnisse vorherzusagen, als Nutzer dies selbst zu tun vermögen. Somit können diese dazu verwendet werden, um Nutzer beim Wiederfinden zu unterstützen. Wie diese Unterstützung in Detail aussehen kann wird in Abschnitt 8 verdeutlicht.

### 7.2.2. Wiederfindensstrategien: *Teleporting* vs. *Orienteering*

In Abschnitt 2.2.4 wurde ausführlich auf die beiden Wiederfindensstrategien *teleporting* und *orienteering* Bezug genommen, welche sich in zahlreichen empirischen Studien als die dominanten Wiederfindensstrategien erwiesen haben. Dort wurde auch ersichtlich, dass man keine allgemeingültige Aussage darüber treffen kann, welche Strategie bei welchem Informationstyp häufiger verwendet wird, oder erfolgreicher ist. In Logstudien zu Re-finding-Verhalten bei E-Mails werden z. B. beide Strategien jeweils als die dominante Strategie beobachtet. Dies lässt den Schluss zu, dass die Erkenntnisse von Kontextfaktoren abhängen und sowohl stark von den Individuen, die als Probanden an der Studie teilnehmen als auch dem jeweiligen E-Mail-Programm, das in der Studie untersucht wird, beeinflusst ist (Dumais et al., 2003; Elswailer, Baillie & Ruthven, 2011; Whittaker et al., 2011). Im Folgenden wird eine der zentralen Fragestellungen der Arbeit, nämlich die Frage nach den verwendeten Wiederfindensstrategien, angesprochen.

Das Balkendiagramm in Abbildung 5.18, welches zur Visualisierung der Zustimmung zur Verwendung bestimmter Wiederfindensstrategien in der Twitter-Umfrage eingesetzt wird, zeigt, dass Probanden überwiegend den Aussagen zur Verwendung von Orienteering-Strategien zustimmen. 59% der Probanden der Twitter-Umfrage stimmen der Aussage zu, Tweets wiederzufinden, indem sie die *Timeline* Abwärts scrollen. 55% stimmen der Aussage zu, das Wiederfindensziel auf der Profilseite des Tweet-Senders zu lokalisieren. Immerhin noch 36% stimmen der Aussage zu, die Favoritenliste zum Wiederfinden zu verwenden. Dem entgegen stehen 36% bzw. 48% der Probanden, die die Nutzung der Twitter-Suche bzw. die Nutzung einer externen Suchmaschine zum Wiederfinden von Tweets ablehnen, sich also tendenziell eher gegen die Verwendung von klassischem *teleporting* aussprechen. Die Twitter-Umfrage zeigt, dass die Probanden eine Präferenz gegenüber dem navigationsbasierten Ansatz äußern. Es ist außerdem zu erkennen, dass es sich mehrheitlich um ein *orienteering* im Sinne eines *opportunistic retrieval* handelt, das kein vorheriges Aufbewahren (*preparatory retrieval*) der Tweets erfordert. Der Verzicht auf das Aufbewahren wurde bereits häufig in anderen Studien festgestellt, da — z. B. bei den KFTF-Studien — Do-nothing-Methoden bevorzugt werden, oder in der Studie von Whittaker et al. in der die Orienteering-Methoden (vgl. Abschnitt 2.2.4 zum Wiederfindensverhalten in unterschiedlichen Domänen) den Vorzug erhielten. Auch im logistischen Regressionsmodell aus Abschnitt 5.7.8.6 erweist sich die Variable *Re-find Method: Search Timeline Person* — also die Suche nach einem Tweet in den gesendeten Tweets eines Nutzers — als signifikanter Prädiktor, der dabei hilft, zu erklären, weshalb Probanden bzw. Nutzer bereits einmal eine frustrierende Wiederfindenserfahrung gemacht haben.

Die Ablehnung der Nutzung von externen Speichern ist in der Umfrage mit 52% sehr groß. Jedoch kann innerhalb der Gruppe der Probanden, die angeben zumindest selten Tweets aufzubewahren, eine signifikant stärkere Zustimmung zur Verwendung der Favoritenliste und externer Speicherorte beim Wiederfinden festgestellt werden (vgl. Abschnitt 5.7.4). Dieser Trend lässt sich auch erkennen, wenn man die Probanden entlang anderer Dimensionen gruppiert wie *Zufriedenheit mit den Wiederfindensoptionen* oder *Schwierigkeit des Wiederfindens* (vgl. Abbildungen 5.21 und 5.20). Nutzer, die zustimmen mehrere Methoden zu verwenden, um Tweets wiederzufinden, charakterisieren Wiederfinden als einfach, bewerten die twitterinternen Features für PIM-Praktiken als ausreichend und waren noch nie frustriert.

Die Eindrücke bzw. die Aussagen der Probanden der Twitter-Umfrage können mithilfe der Ergebnisse der Logstudie objektiv bestätigt werden. Bei der Analyse der Wiederfindensstrategien in Abschnitt 6.4.4.5 (vgl. dort Abbildung 6.24), lässt sich erkennen, dass in kurzen Wiederfindensperioden mit 65% die meisten Re-Klicks im Kontext der eigenen *Timeline* erfolgen und damit Wiederfindensaktionen überwiegend via Scrolling in der *Timeline* gelöst werden. In längeren Wiederfindensperioden finden 79% der Wiederfindensaktionen auf den Profilseiten von Nutzern statt.

Dies entspricht den beiden am stärksten präferierten Wiederfindensmethoden aus der Umfrage, bestätigt also die Präferenz für Orienteering-Verhaltensweisen. Scrolling-Verhalten kann dabei sehr exzessiv Ausmaße annehmen, was anhand der Tweet-Position abgelesen werden kann, die Nutzer erreichen. Das Maximum liegt hier bei 627 ( $Min = 0$ ,  $Mittelwert = 38$ ,  $Median = 7$ ,  $IQR = 41$ ), was einem vergleichsweise langem Abwärts-Scrolling entspricht. Der starke Fokus, der auf Profiseiten der Tweet-Sender liegt, lässt den Schluss zu, dass sich Nutzer gut an diesen erinnern können. Dieser Umstand wird weiter unten noch thematisiert werden. Die Twitter-Suche kommt in beiden Wiederfindensperioden mit insgesamt 3,1% nur selten zum Einsatz, was deutlich macht, dass *teleporting*, also die direkte Schlüsselwortsuche nach einem Tweet, sehr selten zum Wiederfinden verwendet wird. Dies bestätigt die von Teevan et al. formulierte Annahme, dass es sich bei den in ihren Suchanfrageanalysen wiederholten Suchanfragen eher um Monitoring-Verhalten handelt, also dem Beobachten von neuen Tweets zu einem bestimmten Thema (Teevan et al., 2011). Der hohe Anteil (43%) an Re-finding-Suchanfragen, den Elsweiler und Harvey in ihrer Studie festgestellt haben, die das Wiederfinden von Tweets zum Ziel haben, steht stark im Gegensatz zur Häufigkeit der Verwendung der Schlüsselwort-Suche in dieser Studie (Elsweiler & Harvey, 2015).

Zusätzlich kann in der Logstudie auch festgestellt werden, dass Nutzer einen Schwerpunkt auf Wiederfindensmethoden legen, die kein vorheriges Aufbewahren erfordern. Nur in 3,5% (kurze Wiederfindensperiode) bzw. 8% (lange Wiederfindensperiode) der Fälle werden Wiederfindenstrategien genutzt, denen man ein vorheriges Aufbewahren unterstellen kann. Jedoch kann in der Logstudie nicht nachgeprüft werden, ob die Tweets bewusst aufbewahrt wurden, um einen schnelleren Zugriff auf diese zu ermöglichen, oder ob sich Probanden daran erinnern, Tweets favorisiert oder weitergeleitet zu haben und dementsprechend ihre Suchstrategie anpassen. Unterstellt man Probanden jedoch, die Favoriten- und RT-Funktion auch mit der Absicht zu nutzen, Tweets aufzubewahren — wofür es zahlreiche Indizien gibt, die bereits in Abschnitt 7.1 aufgeführt wurden — muss festgestellt werden, dass die Aufbewahrensorte nur selten besucht werden bzw. selten zum Wiederfinden genutzt werden. Bereits in der Twitter-Umfrage wird deutlich, dass Nutzer ihre Favoritenlisten nicht besuchen: 77,4% der Probanden geben an, die Favoritenliste nie oder nur selten zu besuchen. Dies bestätigt sich auch in den Logdaten. Nur 8 von 44 Probanden haben ihre Favoritenliste besucht und nur 0,30% aller Events entfallen auf den Kontext der Favoritenliste.

### **Re-finding-Strategien und Deferred-evaluation-These/O-o-s-o-m-Problem**

Warum werden die eigenen gesendeten Tweets oder die Favoritenliste als Aufbewahrensorte selten besucht? Wenn man versucht, diesen Umstand zu verstehen, bietet die PIM-Literatur (vgl. Abschnitt 2.3) mögliche Erklärungs- bzw. Interpretationsansätze (Jones, Phuwanartnurak et al., 2005; Bergman & Whittaker, 2016). Ein möglicher Ansatz stellt die Deferred-evaluation-These dar. Sie besagt, dass Nutzer die Relevanzprüfung von Informationsobjekten auf einen späteren Zeitpunkt verschieben, da in der aktuellen Situation oft zu wenig Zeit zur Verfügung steht, um Informationen einer vollständigen Relevanzbewertung zu unterziehen, aber dann vergessen, diese Relevanzbewertung tatsächlich durchzuführen und somit nicht auf die Tweets in den Aufbewahrensorten zugreifen. Argumente, die dafür sprechen, dass dies auch auf Twitter der Fall ist, lassen sich bei den Wiederfindensgründen finden, die in der Twitter-Umfrage herausgearbeitet wurden (vgl. Abschnitt 5.8.1). Kode [1.4] *Time Management* beschreibt genau den Umstand, dass Nutzer zu Tweets zurückkehren wollen, um eine Form der *deferred evaluation* durchzuführen, um sich komplett mit dem Tweet und der in ihm enthaltenen Information auseinanderzusetzen. Eine Nutzerbegründung lautet beispielsweise: „I was looking at the link earlier and decided to visit later when I had more time“ [R23]. Auch Probandenkommentare zu Wiederfindensmotiven, bei denen das nachträgliche Speichern als Grund genannt wird — und folglich mit Kode [1.3] *Specific Use* versehen wurden — passen zur Argumentation. Ein Überfluss an Tweets, Zeitbeschränkungen und die daraus resultierende *deferred evaluation* können also mögliche Begründungen für die Beobachtung sein, dass Aufbewahrensorte selten besucht werden. Gegen dieses Argument spricht, dass

Tweets grundsätzlich eher selten favorisiert oder weitergeleitet werden.

Einen weiteren Interpretationsansatz bietet das Out-of-sight-out-of-mind-Problem, welches den Umstand beschreibt, dass Informationsobjekte zwar einer Relevanzbewertung unterzogen werden und anschließend aufbewahrt werden. Jedoch werden die aufbewahrten Tweets nicht besucht bzw. können sich die Probanden nicht erinnern, dass diese Tweets bereits aufbewahrt wurden. In beiden Fällen ist das wesentliche Problem die mangelnde Erinnerungsfunktion — wie sie z. B. Dateien auf dem Desktop oder E-Mails im Posteingang bieten — die von aufbewahrten Tweets letztendlich nicht geleistet werden kann. Die Tweets verlassen die Aufmerksamkeitsebene des Nutzers, so dass keine Erinnerungsfunktion seitens der Objekte selbst besteht und sie einfach vergessen werden.

#### **Re-finding-Pfade, First-Impression-These und Personenfokus**

In den vorherigen Abschnitten wurde evident, dass Nutzer auf das Aufbewahren von Tweets häufig verzichten und auf Orientierung-Strategien vertrauen. Die Präferenz für eine Wiederfindensstrategie, bei der die gesuchten Tweets auf den Nutzerprofilen der Sender lokalisiert werden, deutet darauf hin, dass Nutzer dazu in der Lage sind, den Tweet-Sender zu erinnern. Betrachtet man die Analyse der Re-finding-Pfade aus Abschnitt 6.4.4.6, welche die Wege analysiert, die Nutzer nehmen, um auf die Profileseiten der Tweet-Sender zu gelangen, erkennt man, dass in 30% der Fälle zunächst eine Suche stattfindet, um auf die Profileseite zu gelangen. Man kann dieses Verhalten als suchbasiertes *orienting* bezeichnen. In 70% der Fälle werden einfach die Hyperlinks der *@Mentions* oder Twitter-Namen der gesendeten Tweets des gleichen Nutzers instrumentalisiert, um zum Nutzerprofil zu navigieren. In beiden Fällen werden schließlich Kontextfaktoren zur Orientierung beim Scrolling genutzt oder die Browser-Suche (STRG+F) bemüht, um den Tweet zu finden. Unabhängig von der konkreten Strategie kann man erkennen: der Sender eines Tweets wird erinnert. Betrachtet man in diesem Kontext die von Jones formulierte First-impression-These, gibt diese Aufschluss darüber, worauf Nutzer achten, wenn sie ihre *Timeline* konsumieren (Jones, 2013). Die First-Impression-These besagt, dass Nutzer diejenigen Charakteristika eines Informationsobjekts im Gedächtnis bleiben bzw. zum Wiederfinden verwendet werden, die auch beim initialen Finden genutzt wurden, respektive dazu dienen, eine Priorisierung von Informationsobjekten in einer großen Menge von Informationen vorzunehmen. Da sich Nutzer offensichtlich gut an den Sender der Tweets erinnern können und diesen als Sprungbrett zum Wiederfinden verwenden, stellt dies auch einen Hinweis darauf dar, dass Nutzer schon beim initialen Konsum des Twitter-Streams sehr stark auf den Autor eines Tweets achten. Dies steht auch im Einklang mit der Studie von Counts und Fisher, die feststellten: „factors like type of author impact memory for content (tweets from friends were remembered better than those from organizations)“ (Counts & Fisher, 2011, S.11). Es spricht vieles dafür, dass Nutzer die *Timeline* selektiv wahrnehmen und beim Scannen der neuesten Tweets schon sehr bewusst nach bestimmten Nutzeraccounts Ausschau halten und eine detaillierte Auseinandersetzung mit Tweets entlang bestimmter Tweet-Autoren priorisieren.

Dass Personen oder Accounts in Social-Media-Applikationen eine besondere Rolle spielen, ist nicht besonders überraschend. Dass sie aber auch eine besondere Rolle beim Wiederfindensverhalten einnehmen, wird in der Arbeit evident. Der Vollständigkeit halber werden die Ergebnisse, die die Besonderheit von Personen sowohl im Allgemeinen als auch in Bezug auf das Wiederfinden, deutlich machen, zusammengefasst:

- Favorisierungsmotiv [A2] *Special people as author*: Nutzer favorisieren einen Tweet dann, wenn er von einer besonderen Person (Freunde, Familie, Berühmtheit) geschrieben wurde.
- Das Wiederfinden via Tweet-Sender-Profil ist eine beliebte Wiederfindensstrategie in der Twitter-Umfrage und erweist sich auch in der Logdatenstudie als die zentrale Wiederfindensmethode.
- Im logistischen Regressionsmodell erweisen sich die Features FOLLOWINGCOUNT(LOG), also die Anzahl der Personen, denen ein Nutzer folgt sowie *Re-find Method: Search Timeline Person*,

also das Wiederfinden via *Tweet-Autor-Timeline*, als signifikante Prädiktoren

- Obwohl sich in der qualitativen Analyse der Wiederfindensmotive kein Kode mit Personenbezug herausgebildet hat, werden in zahlreichen Wiederfindensgründen auch Accountnamen erwähnt, bzw. die Suche nach Personen mit im Kommentar angegeben.
- Die Logdatenanalyse zeigt, dass Sitzungen in denen Wiederfindensaktionen stattfinden, mehr Ereignisse im Kontext `USERPROFILEINT` besitzen als normale Sitzungen. Dies ist abermals ein Hinweis darauf, dass beim Wiederfinden häufig Nutzerprofilseiten besucht werden.
- Das Feature *TweetsClickedPerSender*, das das Verhältnis zwischen Rezipient und Tweet-Sendern beschreibt — vorgestellt in Abschnitt 6.4.5.2 und auch weiter oben in Abschnitt 7.2.1 erwähnt — ist eines der nützlichsten Features bei der Vorhersage von Wiederfindenszielen.

#### **Fokus auf mehrfaches Wiederfinden einzelner Tweets**

In Abschnitt 2.3 wurde dargelegt, dass in zahlreichen PIM-Studien ein Fokus auf das mehrfache Wiederfinden einzelner Informationsobjekte festgestellt wurde. Die Twitter-Umfrage liefert diesbezüglich keine näheren Erkenntnisse. In der Logstudie wird jedoch evident, dass 55 der 219 (25%) wiedergefundene Tweets mehr als einmal wiedergefunden wurden. Ein Tweet wurde dreimal erneut geklickt. Interessant ist dabei, dass die Wiederfindensstrategien beim erneuten Wiederfinden konstant bleiben und sich nicht verändern, womit abermals eine der formulierten Leitfragen, nämlich die Frage, ob die Häufigkeit mit der ein Tweet wiedergefunden wird Einfluss auf die Wiederfindensstrategie besitzt, beantwortet wird. Insgesamt kann man nicht von einem Fokus auf einzelne, mehrmals wiedergefundenen Tweets ausgehen.

## **7.3. Zusammenfassung: Identifizierte Probleme**

Das vordergründige Ziel dieser Arbeit liegt nicht in der Konzeption von Lösungsansätzen sondern in der Beobachtung und Beschreibung des Wiederfindensverhaltens von Nutzern auf Twitter. Aus diesen Beobachtungen können jedoch erste Probleme identifiziert und mögliche Lösungen konzipiert werden. Abschließend werden die wesentlichen Erkenntnisse der Diskussion zusammengefasst, so dass auf dieser Basis in Kapitel 8 mögliche Lösungsvorschläge für die Unterstützung von PIM auf Twitter konzipiert und beschrieben werden können.

Folgende Probleme können als zentral identifiziert werden und bilden die Grundlage für Lösungsansätze und Designvorschläge:

- Es wurde festgestellt, dass Nutzer Probleme haben bei der Vorhersage zukünftiger Informationsbedürfnisse. Die Beantwortung der Frage, ob ein Tweet zukünftig von Relevanz sein wird und deswegen wiedergefunden wird, fällt ihnen offenbar schwer. Die Vorhersageexperimente haben gezeigt, dass ein Algorithmus auf der Basis der nötigen Daten dies sehr gut kann. Dieser Umstand sollte für mögliche Lösungsansätze ausgenutzt werden.
- Nutzer haben offensichtlich das Bedürfnis, Tweets aufzubewahren und auch zu diesen zurückzukehren. Hierfür verwenden sie unter anderem das Favoriten-Feature. Es besteht jedoch eine Diskrepanz zwischen Nutzeraussagen und tatsächlichem Verhalten, da die Favoritenliste nur in wenigen Fällen erneut besucht und zum Wiederfinden genutzt wird. Letztendlich gilt es, Lösungen für das *Deferred-evaluation-* bzw. *Out-of-sight-out-of-mind-*Problem zu finden.
- Nutzer verwenden fast ausschließlich navigationsbasierte Wiederfindensstrategien. Jedoch gibt es auch Anzeichen dafür, dass Nutzer *teleporting*, also suchbasierte Wiederfindensstrategien, verwenden würden, falls dieses Verhalten seitens Twitter besser unterstützt werden würde. Auch im Kontext des navigationsbasierten Wiederfindens besteht Raum für Verbesse-

rungen, in dem dieses so unterstützt wird, dass weniger Zeit- und Interaktionsaufwand nötig ist.

Für diese Probleme sollen im Folgenden Lösungsansätze konzipiert und diskutiert werden.



## Wiederfindensverhalten unterstützen: Lösungsansätze und Designvorschläge

Die Lösungsansätze und Designvorschläge werden vor allem in Bezug auf Twitter als Anwendung konzipiert, da sie somit sowohl im Web-Client als auch in den mobilen Applikationen genutzt werden können. Für prototypische Implementierungen können auch eigene Twitter-Clients oder Browser-Erweiterungen entwickelt werden, die mit den nötigen Lösungsansätzen ausgestattet werden. Prinzipiell kann auch die für die Sammlung der Logdaten genutzte Chrome-Erweiterung so weit verbessert und erweitert werden, so dass sie die angeführten Lösungsvorschläge umsetzt. Darauf aufbauend können neue Studien konzipiert werden, welche die vorgestellten Designvorschläge evaluieren.

Designansätze für die Unterstützung des Wiederfindensverhaltens bewegen sich in zwei Bereichen: Algorithmische Ansätze für das *Re-ranking* von Tweets in der *Timeline* oder der SERP und SUI-basierte Ansätze. Eine Möglichkeit, um clientseitige Veränderungen vorzunehmen und somit sowohl algorithmische als auch SUI-Verbesserungen zu ermöglichen, bieten Browsererweiterungen, welche im akademischen Kontext weit verbreitet sind. Im Kontext der Informationssuche im Web bzw. mit Websuchmaschinen werden auch Browsererweiterungen erstellt, die mögliche Lösungsansätze für die Unterstützung von Re-finding und Wiederbesuch von Websites umsetzen. Morris und Horvitz entwickeln das  $S^3$ -System, das es dem Nutzer erlaubt, Suchanfragen und relevante Ergebnisse für diese Suchanfragen zu speichern, um diese wiederzufinden (M. R. Morris & Horvitz, 2007). Vergleichbare Ansätze werden bereits bei Komlodi, Soergel und Marchionini getestet, die Experimente mit Suchverläufen — UI-Bereiche, die Suchanfrage-Ergebnis-Kombinationen visualisieren — für Informationssysteme im Rechtskontext durchführen (Komlodi et al., 2006). Die Erweiterung *SearchBar* von Morris, Ringel Morris und Venolia ermöglicht das Durchsuchen und Verwalten von getätigten Suchanfragen und geklickten Suchergebnissen (D. Morris et al., 2008). Die *Re:search Engine* von Teevan kombiniert neue und alte Suchergebnisse in einer SERP, um sowohl das Auffinden neuer Information als auch Wiederfinden zu unterstützen (Teevan, 2007). Das SUPRA-Framework von Kawase et al. ist eine Bibliothek zur kontextbasierten Echtzeitvorhersage von Navigationsaktivitäten des Nutzers, die ihn beim Wiederbesuch von Seiten unterstützen soll, indem es ihm Vorschläge unterbreitet (Kawase et al., 2011). Jhaveri und Räihä verwenden in ihrer Browsererweiterung *Thumbnails*, um das Wiedererkennen und somit das Wiederfinden von bereits besuchten Seiten zu erleichtern (Jhaveri & Räihä, 2005).

Die exemplarisch ausgewählten Forschungsarbeiten zu Erweiterungen, die Re-finding-Unterstützung bieten, sollen aufzeigen, welche unterschiedlichen Varianten von Unterstützungsmöglichkeiten existieren. Es ist anzumerken, dass die Designvorschläge und Lösungsansätze die Tatsache außer Acht lassen, dass externe Speicher (z. B. Bookmarking-Dienste wie *Delicious*<sup>1</sup>, *Diigo*<sup>2</sup> oder *Pocket*<sup>3</sup>) genutzt werden, um Tweets zu speichern, zu verwalten und wiederzufinden. Ein vergleichbares Hilfsmittel stellt das Tool *Packrati.us*<sup>4</sup> dar, welches URLs aus eigenen gesendeten Tweets automatisch zu Bookmarking-Diensten hinzufügt.

Folgende Fragen werden in den nächsten Abschnitten angesprochen:

- Was passiert, wenn Nutzer wichtige Tweets vergessen, wie wollen und können sie an diese erinnert werden?
- Wie kann Wiederfinden sowohl vorab als auch während des Wiederfindensprozesses unterstützt werden? Wie können vermeintlich seltener genutzte Wiederfindensstrategien wie *Teleporting* verbessert werden?
- Wie kann man die Aufmerksamkeit des Nutzers auf Tweets lenken, die sie vermutlich wiederfinden wollen?

Diese Fragen werden anhand der in Abschnitt 7.3 identifizierten Probleme diskutiert, wobei auch der Ist-Zustand bestimmter Twitter-Funktionen vorgestellt wird. Wichtig ist, dass diese Fragen bzw. die Lösungsansätze und Designvorschläge auf einer Definition von Re-finding aufbauen, wie es — zumindest in der Logstudie — in Form des Proxies definiert ist, nämlich als Across-Session-Re-klick. Für eine Umsetzung müsste also ein *click caching* sämtlicher Nutzerklicks erfolgen.

## 8.1. Erinnerungsfunktion für aufbewahrte Tweets

In der Arbeit wurde eine Diskrepanz festgestellt zwischen der Häufigkeit mit der Probanden angeben, die Favoriten-Funktion zum Aufbewahren von Tweets zu nutzen und der Häufigkeit, mit der die Favoritenliste zum erneuten Besuch dieser Tweets aufgesucht wird. Nur bei 5% aller Wiederfindensaktionen wurde die Favoritenliste genutzt. Zwei mögliche Interpretationen dieser Beobachtung wurden entlang der Deferred-evaluation-Hypothese und dem Out-of-sight-out-of-mind-Problem formuliert. Das Deferred-evaluation-Problem besagt, dass in der aktuellen Situation nicht genügend Zeit zur Verfügung steht, um die Relevanz eines Tweets vollständig zu bewerten und diese Bewertung auf einen späteren Zeitpunkt verschoben wird. Hierfür wird dieser Tweet zunächst aufbewahrt. Die Out-of-sight-out-of-mind-These besagt, dass Nutzer einen Tweet als relevant erachten und ihn für einen späteren Zeitpunkt aufbewahren, aber den Aufbewahrungsort — in diesem Fall die Favoritenliste — nicht zum Wiederfinden des Tweets nutzen. Beiden Fällen liegt ein Erinnerungsproblem zu Grunde. Das Aufbewahren eines Tweets in einen nur mittelbar sichtbaren und zugänglichen Speicherort lässt Nutzer vergessen, diese Handlung durchgeführt zu haben. Mögliche Erinnerungsfunktionen, wie es E-Mails im Posteingang oder Dateien auf dem Desktop zu leisten vermögen, können von Tweets so nicht geleistet werden. Eine mögliche Lösung für dieses Problem stellt die Einführung einer Erinnerungsfunktion dar. Twitter bietet dem Nutzer über dessen Profil eine detaillierte Auswahl für Einstellungsmöglichkeiten zu E-Mail- und Webbenachrichtigungen. Webbenachrichtigungen erscheinen z. B. als Pop-up am rechten Rand des Bildschirms, wenn ein eigener gesendeter Tweet weitergeleitet oder favorisiert wird, oder wenn der Nutzer eine Direktnachricht erhält.

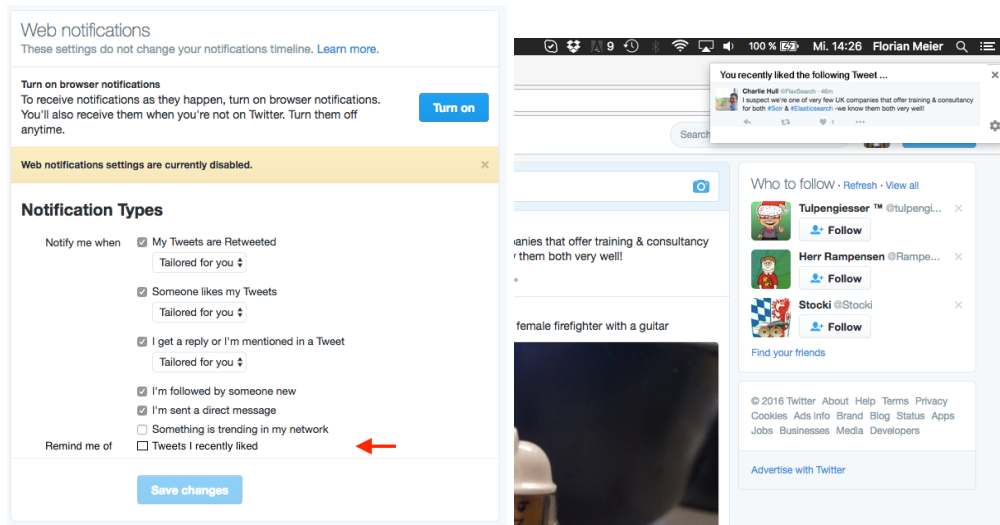
---

<sup>1</sup><http://del.icio.us>

<sup>2</sup><http://www.diigo.com>

<sup>3</sup><http://www.getpocket.com>

<sup>4</sup><http://packrati.us>



(a) Web-Notifications-Option, für die Erinnerungsfunktion an favorisierte Tweets (b) Beispiel eines Erinnerungs-Pop-ups für einen favorisierten Tweet

**Abbildung 8.1.:** Visualisierung von Konfiguration und Anzeige der Webbenachrichtigung favorisierter Tweets (Modifikation der Screenshots von: [https://twitter.com/settings/web\\_notifications](https://twitter.com/settings/web_notifications) bzw. <https://twitter.com>)

Abbildung 8.1a zeigt, wie die Wahlmöglichkeit der Erinnerungsfunktion in die Einstellungen eingebettet werden könnte. Abbildung 8.1b zeigt exemplarisch, wie ein solches Erinnerungs-Pop-up visualisiert werden könnte. Natürlich sollte nicht für jeden favorisierten Tweet eine solche Erinnerung erfolgen. Um zu vermeiden, dass Nutzer durch zu häufige Erinnerungs-Pop-ups gestört werden, könnten nur diejenigen Favoriten angezeigt werden, die durch einen Algorithmus als Bookmark-Favoriten vorhergesagt werden. Gorrell und Bontechva widmen sich in ihrer Arbeit der Frage, ob ein Favorit mit der Absicht erstellt wurde, um ein Lesezeichen zu setzen (Gorrell & Bontcheva, 2016). Eine einfache Heuristik könnte z. B. sein, dass favorisierte Tweets, die URLs enthalten, mit der Absicht favorisiert wurden, um die URL zu einem späteren Zeitpunkt wiederzufinden. Zudem müssten auch Erinnerungsintervalle wählbar sein. Optimale Erinnerungsintervalle sollten durch zusätzliche Experimente/Nutzerstudien bestimmt werden. Die Erinnerungsfunktion könnte nicht nur als Webbenachrichtigung umgesetzt werden. Eine Realisierung via E-Mailbenachrichtigungen wäre ebenso denkbar und vermeintlich weniger invasiv.

#### Entkoppelung von *like* und *bookmark*

Eine weitere sinnvolle Veränderung könnte in der Entkopplung der Like- und Lesezeichen-Funktionalität bestehen. Die Analyse und Diskussion der Favorisierungsgründe in Abschnitt 5.6.3 hat gezeigt, dass die Nutzung der Funktion zwei Dimensionen besitzt, wobei *Like* und *Bookmark* die beiden dominanten Vertreter dieser Dimensionen darstellen. Diese beiden Dimensionen sind zwar konzeptuell und in Bezug auf das mentale Modell der Nutzer voneinander getrennt, jedoch in einem UI-Feature (dem Stern/dem Herz) miteinander verknüpft. Man stelle sich eine hypothetische Twitter-Sitzung vor, in der vom Nutzer mehrere Tweets favorisiert werden, wobei nur eine Favorisierung dazu gedacht ist, den Tweet aufzubewahren und die anderen als klassische *Likes* intendiert sind. Diese Situation trägt sicherlich dazu bei, dass das Erinnern des Aufbewahrens noch schwerer fällt. Eine Trennung der Like- und Bookmark-Funktion in zwei separate UI-Features könnte zu einer Lösung führen, bei der Nutzer sich auch leichter daran erinnern, dass sie einen Tweet aufbewahrt haben.

## 8.2. Re-finding-Verhalten unterstützen

Wiederfinden könnte sowohl prophylaktisch unterstützt werden, also bevor der Nutzer das Bedürfnis nach Wiederfinden noch gar nicht geäußert hat, oder zu Beginn, also während des Wiederfindensprozesses. Im Folgenden werden vier Varianten vorgestellt.

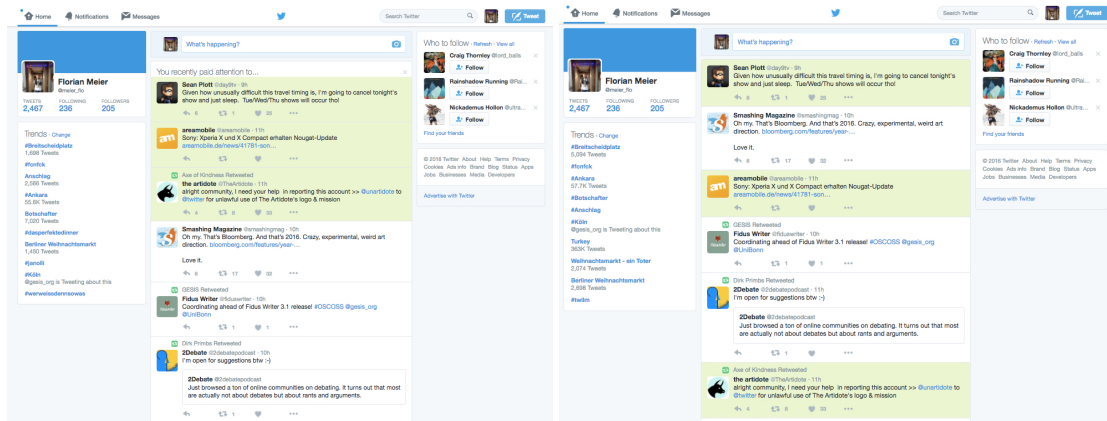
### 8.2.1. *Interleaving* von Timeline-Tweets mit Wiederfindenszielen

Eine latente und dadurch prohibitive — also bereits bevor beim Nutzer die Erkenntnis reift, dass der gesehen Tweet relevant gewesen wäre und ein Wiederfindensprozess angestoßen wird — Möglichkeit, Wiederfinden zu unterstützen, stellt das *interleaving* also das Verschachteln von neuen und älteren, wiederzufindenden Tweets dar. Diese Idee basiert auf einer Vorgehensweise wie sie von Teevan in der Browsererweiterung *Re:search Engine* für Websuchmaschinenergebnisse umgesetzt und getestet wurde (Teevan, 2007). Teevan argumentiert, dass Nutzer häufig das Bedürfnis haben, sowohl neue Informationen zu bekommen als auch innerhalb derselben Session bei der gleichen Suchanfrage alte Ergebnisse wiederfinden wollen. Der *Re:search*-Ansatz versucht genau dies zu unterstützen, indem ältere Suchergebnisse mit neuen angereichert bzw. verschachtelt werden (Teevan, 2007). Diese Idee könnte auch für Twitter umgesetzt werden bzw. wird dies schon in Ansätzen für ältere, noch nicht betrachtete Tweets mithilfe der *While-you-were-away*-Funktion<sup>5</sup>, realisiert. Diese Funktion zielt auf eine kuratierte *Timeline* ab, die die unmittelbar nach dem letzten Besuch gesendeten Tweets in einem leicht veränderten Ranking vor dem klassischen *Stream*, der die aktuellsten Tweets umgekehrt-chronologisch zeigt, voranstellt. Ähnlich kann dies bei Wiederfindenszielen realisiert werden. Ältere Tweets aus früheren Sessions für die vorhergesagt wird, dass sie wiedergefunden werden, könnten in den *Stream* der neuesten Tweets eingebettet werden. Dies befreit den Nutzer davon, Vorhersagen zu zukünftigen Informationsbedürfnissen zu treffen, da ein automatisches Verfahren, das diese Vorhersagen ohnehin besser beherrscht als der Nutzer selbst, diese Aufgabe für ihn übernimmt. Diese Methode könnte auch auf die in Abschnitt 8.1 beschriebenen Bookmark-Tweets angewendet werden. Statt einer Erinnerung via E-Mail- oder Webbenachrichtigung könnten diese Tweets in die *Timeline* zwischen neuen Tweets eingebettet werden, so dass ein Wiedererkennungseffekt entsteht.

Die Abbildungen 8.2a und 8.2b visualisieren zwei mögliche Varianten dieses Lösungsansatzes. In beiden Varianten sind die potentiellen Wiederfindensziele grün hervorgehoben. Die Hervorhebung dient nur der Veranschaulichung und wäre in einer tatsächlichen Umsetzung nicht vorhanden. Der erste Ansatz in Abbildung 8.2a ist vergleichbar mit dem *In-case-you-missed-it*-Feature und zeigt potentielle Wiederfindensziele im Block am Anfang der *Timeline*. Mit einem Hinweissatz wie *You recently paid attention to...* kann der Nutzer bewusst auf diese Wiederfindensziele aufmerksam gemacht werden. Der zweite Ansatz in Abbildung 8.2b ist weniger offensiv und bettet Wiederfindensziele in die neueren Tweets ein. Hierfür wären verschiedene *Interleaving*- oder *Merge*-Strategien denkbar. So könnte die Zeit, die seit dem ersten Klick vergangen ist, als Verhältnis genutzt werden, um die Position in der zusammengefassten *Timeline* zu bestimmen. Auch hier wären zusätzliche Fragen zu klären:

- Wie verträgt sich diese Funktion mit dem *In-case-you-missed-it*-Feature?
- Wie viele Tweets sollten bei der Darstellung im Block angeboten bzw. in die *Timeline* eingebettet werden?
- Wie alt können Tweets sein, die dem Nutzer erneut angeboten werden? Eine mögliche Lösung

<sup>5</sup>Kürzlich in *In case you missed it...* umbenannt.



(a) Verflechtungsmethode 1: Wiederfindensziele im Block (b) Verflechtungsmethode 2: Wiederfindensziele in Timeline integriert

**Abbildung 8.2.:** Vergleich zweier Interleaving-Methoden. Grün hervorgehobene Tweets sind ältere, als potentielle Wiederfindensziele identifizierte Tweets, die in den aktuellen *Stream* integriert werden. (Modifikation eines Screenshots von: <https://twitter.com>)

bestünde darin, eine separate, gerankte Liste an Wiederfindenszielen zu erstellen, wobei immer ein gewisser Teil dieser Liste eingebettet wird (z. B. die Top fünf) und Tweets aus dieser Liste gelöscht werden, wenn sie wiedergefunden wurden.

Dies sind Fragestellungen, die in separaten Studien geklärt werden müssten. Zudem muss erwähnt werden, dass ein Eingriff in die *Timeline* eine Veränderung darstellt, die zu Orientierungsproblemen bei den Nutzern führen kann, wenn die klassische Chronologie der *Timeline* intransparent wird.

### 8.2.2. Verlauf geklickter Tweets

Die anfangs vorgestellten Browser-Erweiterungen zur Unterstützung des Wiederfindensverhaltens bei der Websuche besitzen eine Gemeinsamkeit, die automatisch zu einem Konzept führt, nämlich einer Verlaufshistorie geklickter oder interagierter Webdokumente. Auch im Kontext von Twitter könnte eine separate Liste an Tweets, mit denen man interagiert hat, Wiederfinden unterstützen. Die Erstellung eines Verlaufs geklickter Tweets in Form einer separat besuchbaren *Timeline* stellt eine einfache Variante dar, die den Wiederbesuch bzw. das Wiederfinden von Tweets unterstützen würde. Es ist jedoch fraglich, ob die Einführung einer solchen zusätzlichen Seite, respektive *Timeline* — die letztendlich nur einen Zweck erfüllt — sich so einfach in das Gesamtkonzept der Social-Media-Anwendung integrieren ließe. Auch die reine Existenz dieser separaten *Timeline* stellt wahrscheinlich noch keine endgültige Lösung dar, da diese auch in erster Linie besucht werden müsste, was abermals zu einem Problem werden kann, wenn die nötige Erinnerungsfunktion bzw. der Anreiz für den Besuch der separaten *Timeline* der geklickten Tweets fehlt.

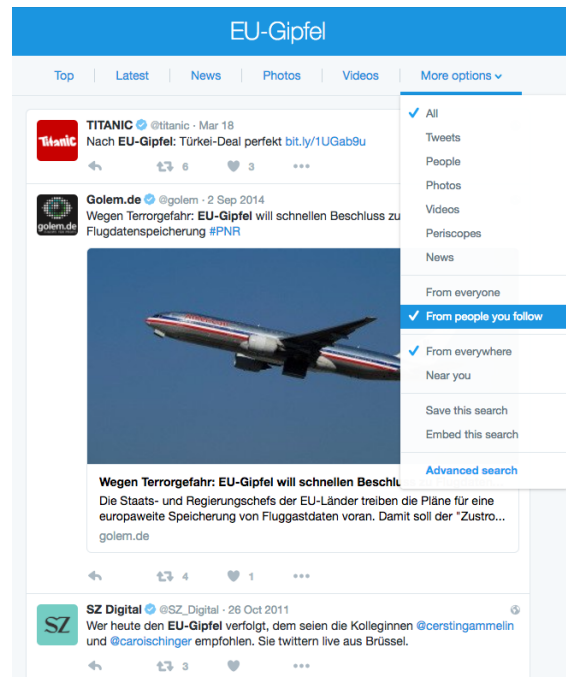
### 8.2.3. Unterstützung des Teleporting-Verhaltens: *Learning to re-find*

Die direkte Suche nach Wiederfindenszielen via Twitter-Suche wurde insgesamt nur in 3,1% der Fälle genutzt. Dumais et al. argumentieren, dass die Suche nach alten Tweets durch die Twitter-Suche sehr schlecht unterstützt ist (Dumais et al., 2014). Hingegen konnte bei der Studie von

Elsweiler und Harvey festgestellt werden, dass sehr viele Nutzer Re-finding-Aktionen über die Twitter-Suche zu lösen versuchen (Elsweiler & Harvey, 2015). Die Twitter-Suche unterstützt die Verwendung des `@Screennames` (z. B. `@meier_flo`) als Operator, um die gesendeten Tweets dieses Nutzers zu durchsuchen. Dies funktioniert sowohl in Bezug auf andere Nutzer als auch auf den eigenen Account, also die eigenen gesendeten Tweets. Probeweise Versuche der Nutzung dieses Operators zeigen, dass die Suche hiermit sehr gut funktioniert und Wiederfinden somit durchaus erleichtert wird. Interessanterweise können in den Suchanfragen der Logstudie elf Suchanfragen identifiziert werden, die einen `@Screenname` und einen weiteren Queryterm (z. B. `@delsweil kftf`) enthalten. Ein Grund, warum dies eher selten auftritt, kann vermutlich darauf zurückgeführt werden, dass Nutzer diesen Operator nicht kennen. Ein weiterer Grund besteht sicherlich in der Tatsache, dass man bei einer Suchanfrage, die mit einem `@Screenname` beginnt, die Autocomplete- bzw. Autosuggest-Funktion der Suche sofort mögliche Nutzer vorschlägt und bei einer Bestätigung der Suchanfragen, durch einen Klick oder dem Drücken der Enter-Taste, dieser Nutzer sofort auf die Profilseite geleitet wird. Es liegt also ein gewisser *bias* vor, der Nutzer unmittelbar die Profilseite des Nutzers besuchen lässt und ihn davon abhält, einen zusätzlichen Queryterm einzugeben. Die Nutzung des Operators würde also eine mögliche Lösung darstellen, besitzt jedoch offensichtlich Schwächen. Hinzu kommt die Frage: Was passiert, wenn sich Nutzer nicht an den Sender eines Tweets erinnern können und somit auf die Suche mit einem inhaltlichen Term angewiesen sind? Es bestünde auch die Möglichkeit, bei einer herkömmlichen Suchanfrage unter dem Reiter *more options* die Suche auf Accounts einzuschränken, denen man selber folgt, was schließlich ein Wiederfinden ebenfalls erleichtern würde. Abbildung 8.3 zeigt die Einstellungsoption für eine Suche in den gesendeten Tweets aller Followees. Selbst wenn das Bewusstsein über die Existenz dieser Optionen vorhanden ist: Statistiken zur Websuche zeigen, dass nur ein Bruchteil der Nutzer Operatoren bei der Suche verwendet (Jansen & Spink, 2006). In der von Teevan, Ramage und Morris durchgeführten Twitter-Querylogstudie werden in 3,4% aller Suchanfragen `@Mentions` verwendet (Teevan et al., 2011, S.38).

Die Twitter-Suche ist auf Echtzeitsuche und *Monitoring* optimiert. Die Frage, die sich im vorliegenden Kontext stellt ist: Wie könnte eine Schlüsselwortsuche auf Wiederfinden hin optimiert werden? Möglichkeiten bieten sogenannte Learning-to-Rank-Verfahren, wie sie sich bei der Erstellung von SERPs in der Websuche bereits etabliert haben und immer mehr auch im Bereich Personal-Search-Anwendung finden (Tyler et al., 2010; Carmel et al., 2015; X. Wang et al., 2016). Zu Personal-Search-Szenarien gehören E-Mail-Suche, Desktop-Suche, Suche auf mobilen Geräten oder auch Suche auf Social-Media-Plattformen wie Twitter. Wang et al. gehen der Frage nach, wie Klick-Modelle im Personal-Search-Kontext genutzt werden können, da in diesem Kontext Klick-Daten eine hohe Spärlichkeit aufweisen (X. Wang et al., 2016). Diese Spärlichkeit trifft auch auf Twitter zu, da auch dort, wie auch die vorliegende Arbeit zeigen kann, nur wenige Tweets tatsächlich geklickt werden (vgl. Tabelle 6.3). Die Suche in Twitter stellt aber wohl auch deshalb eine Sonderrolle dar, da neben den persönlichen Klick-Modellen, die folglich einen gewissen Auswahl-Bias in ihren Berechnungen der Klick-Wahrscheinlichkeit in Betracht ziehen müssen, was von Wang und Kollegen entsprechend berücksichtigt wird, auch auf bestehende Klick-Modelle zurückgegriffen werden kann, die die Klick-Wahrscheinlichkeiten aus einer riesigen Klick-Menge der ganzen Plattform ableiten können. Ohne im Detail auf diese Verfahren einzugehen, wäre eine Optimierung des Tweet-Rankings bei der Suche anhand von Features unter Verwendung eines Learning-to-Rank-Verfahrens möglich. Als Merkmale können die in Abschnitt 6.4.5.2 vorgestellten Features genutzt werden, die schon bei den Vorhersageexperimenten genutzt wurden. Erste Studien können hier z. B. auf der Basis des LTR-Plug-Ins von Apache Solr<sup>6</sup> einfache Baseline-Rankings basierend auf Methoden wie  $TF*IDF$  oder  $BM25$  mit einem gelernten Ranking vergleichen.

<sup>6</sup><https://github.com/bloomberg/lucene-solr/tree/master-ltr-plugin-release/solr/contrib/ltr>



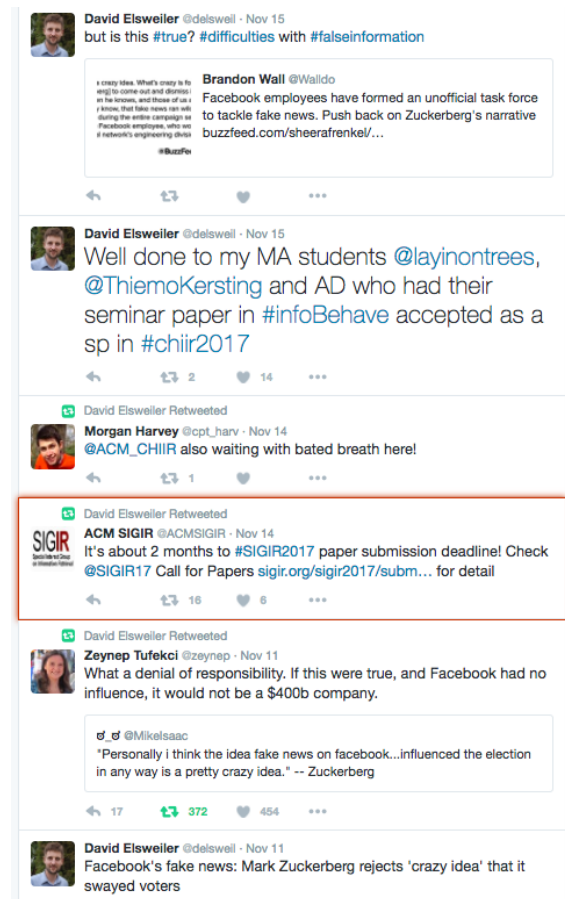
**Abbildung 8.3.:** Erweiterte Suchfunktion der Twitter-Suche, die das Durchsuchen der gesendeten Tweets der Followees erlaubt

(Screenshot von: <https://twitter.com/search-home>)

#### 8.2.4. Unterstützung des Orienteering-Verhaltens

Die navigationsbasierte Suche, bei der Nutzer die wiederzufindenden Tweets auf den Profilen der Tweet-Sender lokalisieren, wurde als die dominante Wiederfindensstrategie für lange Wiederfindensperioden identifiziert. Für kurze Wiederfindensperioden wurde das Abwärts-Scrolling auf der *Timeline* als die dominante Strategie identifiziert. In beiden Fällen erfolgt das Lokalisieren des Wiederfindensziels durch Navigation, Scrolling oder auch unter Verwendung der Browser-Suche (Tastaturkürzel: STRG/CMD+F). Durch die Anordnung der Tweets in umgekehrt-chronologischer Reihenfolge — unabhängig, ob es sich um die eigene *Timeline* oder um die gesendeten Tweets eines Nutzers handelt — fehlen dem Nutzer gewisse Kontextfaktoren bei der Suche. Beim Scrolling-Prozess orientiert sich der Nutzer vermutlich anhand verschiedener Kontextfaktoren wie anderen Tweets, die ihm während des initialen Auffindens begegnet sind, was dem Rang bei der Websuche entsprechen würde, der als Faktor dazu beiträgt, dass Nutzer sich an ein Websuchmaschinenergebnis erinnern können (Teevan, 2007). Um schließlich den Nutzer bei diesem Annäherungsprozess zum gesuchten Tweet zu unterstützen und einen möglichen Moment, bei dem er das Wiederfindensziel erkennt, zu beschleunigen, kann eine optische Hervorhebung dieses Wiederfindensziels erfolgen. Das optische Hervorheben von Tweets, die potentielle Wiederfindensziele darstellen, könnte immer dann erfolgen, wenn erkannt wird, dass der Nutzer sich in einer Wiederfindenssituation befindet. Die in der Arbeit vorgeschlagene Methode zur Vorhersage von Wiederfindenszielen könnte dann genutzt werden, um die entsprechenden Tweets optisch hervorzuheben.

Abbildung 8.4 zeigt, wie eine optische Hervorhebung aussehen könnte. Diese Hervorhebung würde das Wiedererkennen und Wiederfinden eines Tweets potentiell beschleunigen und dem Nutzer mehrere Auf- und Abwärts-Scrollvorgänge sowie einen vermeintlich frustrierenden Wiederfindensprozess ersparen. Vorteil dieser Methode ist, dass sie unabhängig von der tatsächlichen Wiederfin-



**Abbildung 8.4.:** Exemplarische Darstellung eines hervorgehobenen Wiederfindensziels (Modifikation eines Screenshots von: <https://twitter.com>)

denstrategie eingesetzt werden könnte. So können Wiederfindensziele auf der eigenen *Timeline*, oder auf den Profiseiten anderer Nutzer genauso hervorgehoben werden wie auch bei einem SERP-Ergebnis einer Suchanfrage, die als Wiederfindenssuchanfrage identifiziert wird. Dies erlaubt eine Kombination mit dem Vorschlag einer Verbesserung des Rankings aus Abschnitt 8.2.3. Auch eine Kombination mit dem Vorschlag der Verschachtelung von alten und neuen Tweets, um Wiederfinden zu erleichtern, könnte von dieser Methode profitieren. Wie auch bei dem Lösungsvorschlag der Erinnerungsfunktion kann dem Nutzer letztendlich die Wahl gelassen werden, in welchen Kontexten er eine Hervorhebung erlauben würde und mit welcher Häufigkeit Hervorhebungen vom System gemacht werden sollen. Wie bei allen Designvorschlägen gilt auch in diesem Kontext, dass die Notwendigkeit nach zusätzlichen Nutzerstudien besteht. Beispielsweise steht die Frage nach der maximalen Anzahl an hervorzuhebenden Tweets im Raum, wenn mehrere potentielle Wiederfindensziele existieren, die noch zu einem guten Nutzungserlebnis beitragen, ohne den Nutzer zu stören. Ebenfalls wäre die Frage nach dem zeitlichen Horizont, also wann mit der Hervorhebung eingesetzt werden sollte (nach 5 Sekunden, nach 15 Sekunden etc.), experimentell zu klären.



## 8.3. Fazit zu den Designvorschlägen

Die letzten Abschnitte haben zahlreiche Lösungsansätze bzw. Designvorschläge vorgestellt, die aufzeigen, wie Wiederfindensverhalten in unterschiedlichen Phasen und bei unterschiedlichen Wiederfindensstrategien unterstützt werden könnte. Jedoch ist zu beachten, dass diese nicht so ohne Weiteres in Twitter implementiert werden können bzw. deren Umsetzung automatisch eine Verbesserung der Wiederfindenserfahrung bedeutet. Folgende Punkte sollten beachtet werden:

- Einige der Verfahren sind zwar miteinander kombinierbar, jedoch müssten die einzelnen Verfahren erst alleine und dann in Kombination getestet werden. Vorzugsweise würden die Verfahren in mehreren A/B-Tests ergänzt durch die Experience-Sampling-Methode miteinander kontrastiert werden. Dies könnte jedoch — wie die Erhebung der Logdaten — über eine Erweiterung für den Chrome Browser realisiert werden. Die Designvorschläge sind gleichzeitig als Ausblick für zukünftige Studien zu interpretieren.
- Den Vorschlägen liegt eine Re-finding-Definition zu Grunde, die einen erneuten Klick vorsieht. Es könnten zwar auch andere Proxies verwendet werden, jedoch würde die Anzahl der potentiellen Wiederfindensziele stark steigen. Wie damit umzugehen ist, müsste ebenfalls in Studien ermittelt werden.

Abschließend werden nochmals die wichtigsten Erkenntnisse der Arbeit zusammengefasst sowie ein Ausblick über mögliche weitere Forschungsansätze gegeben.



# 9

## Fazit und Ausblick

Im finalen Kapitel der Arbeit werden die Hauptbeiträge der Studien zusammengefasst. Es erfolgen kritische Anmerkungen zur eingesetzten Methodik sowie zur Analyse von Social-Media-Plattformen im Allgemeinen. Final werden offene Fragen formuliert und Anknüpfungspunkte für zukünftige Forschungsarbeiten präsentiert.

### 9.1. Hauptbeiträge der Arbeit in Zusammenfassung

In der Einleitung der Arbeit wurde die Beobachtung, Dokumentation, Beschreibung und Interpretation von Wiederfindensverhalten auf Twitter sowie die Erarbeitung von Lösungskonzepten zur Unterstützung des Wiederfindens zum Hauptziel der Arbeit erklärt. Im Folgenden werden die Erkenntnisse, welche dazu beitragen, dieses Ziel zu realisieren, vor dem Hintergrund der in der Einleitung formulierten Beobachtungen, die die Arbeit motiviert haben, dargestellt. Eine detaillierte Darstellung und Diskussion der Ergebnisse erfolgte bereits im Zuge der Meta-Inferenz.

#### **Quantifizierung der PIM-Verhaltensweisen auf Twitter**

Die Arbeit nimmt eine erste Quantifizierung von PIM-Praktiken im Social-Media-Kontext vor. Es zeigt sich, dass auch in diesem Kontext die Verwaltung und Rückkehr zu Informationen ein unerwartet häufiges Nutzerbedürfnis darstellt, das bis zu drei Viertel der Probanden betrifft. Bezieht man dies auf die in der Einleitung zitierte Hypothese von Bergman und Whittaker, die PIM-Probleme als tief-verwurzelt im menschlichen Verhalten und losgelöst von technologischen Fortschritten oder Daten charakterisiert, so kann man diesen Aussagen, zieht man die Erkenntnisse der Arbeit in Betracht, uneingeschränkt zustimmen (Bergman & Whittaker, 2016).

#### **Kategoriensystem von Wiederfindensgründen**

Die Arbeit nimmt eine systematische Auswertung von Wiederfindensmotiven vor. Die Motive, die Nutzer dazu bewegen Tweets wiederzufinden, können sehr unterschiedlich sein. Sie reichen von Wiederfinden einfacher Zitate, Links oder Bilder, über Wiederfinden, um bestimmte Handlungen mit dem Tweet vorzunehmen, wie ihn mit anderen Personen zu teilen, bis hin zu sentimental Motiven, wie das Betrachten eigener gesendeter Tweets aus Nostalgie. Bemerkenswert ist die Tatsache, dass Wiederfindensbedürfnisse existieren, die nicht durch einzelne Tweets befriedigt werden können, sondern bei denen eine Vielzahl an Tweets betrachtet werden. Häufig tritt dies in Zusammenhang mit eigenen gesendeten Tweets auf, um beispielsweise sämtliche Bilder, die man

gesendet hat, erneut zu betrachten und dadurch in Erinnerungen zu schwelgen. Insgesamt konnten 15 unterschiedliche Beweggründe festgestellt werden.

#### **Modell zur Vorhersage von frustrierenden Wiederfindenssituationen**

Wiederfinden kann Probleme bereiten und Nutzer können bei der Suche nach Informationsobjekten frustriert zurückgelassen werden. Dies zeigt sich — in zahlreichen anderen Studien, die in Abschnitt 2.2.4 vorgestellt werden — auch in dieser Arbeit in Bezug auf Tweets. Dies offenbart sich sowohl im logistischen Regressionsmodell, welches die Faktoren, die Wiederfinden zu einer frustrierenden Erfahrung werden lassen, ergründet als auch bei der Auswertung der Logdaten, die zeigen, dass Sitzungen in denen Wiederfinden stattfindet, im Vergleich zu normalen Twitter-Sitzungen weitaus komplexer sind und mehr Arbeit für den Nutzer darstellen. Interessant ist, dass die Variable FOLLOWINGCOUNT(LOG), also die Anzahl der Followees eines Nutzers, signifikant zur Vorhersage von frustrierenden Wiederfindenserfahrungen beiträgt. Dies steht im Einklang mit der in der Einleitung beschriebenen Beobachtung von Gomez-Rodriguez et al., die feststellen, dass ab einer Followee-Zahl von etwa 100 Personen, Twitter-Nutzer unter einer Informationsflut leiden (Gomez-Rodriguez et al., 2014).

#### **Kategoriensystem von Motiven zur Nutzung der Favoriten-Funktion**

Auf der Basis von Probandenkommentaren zu Beweggründen und Motiven der Favoriten-Funktion wurde mithilfe qualitativer Datenauswertung ein Kodierschema entwickelt, das 25 verschiedene Gründe aufzeigt, die Favoriten-Funktion zu nutzen. Neben dem klassischen *like* wurde Favorisieren, um ein Lesezeichen auf Tweets zu setzen — also diese in der Favoritenliste aufzubewahren und den Zugang zu ihnen zu erleichtern — als Hauptmotivation für die Nutzung der Funktion aufgedeckt. Auch die quantitativen Daten der Twitter-Umfrage konnten zeigen, dass Nutzer die Favoriten-Funktion sowie die RT-Funktion zum Aufbewahren von Tweets verwenden. Zudem wurde aufgedeckt, dass Nutzer die Tweets aufbewahren, auch signifikant stärker den Aussagen zustimmen, diese Speicherorte zum Wiederfinden zu verwenden. Hier kann jedoch eine Diskrepanz zwischen Nutzeraussagen und tatsächlichem Verhalten, welches via Twitter-Logstudie aufgezeichnet wurde, festgestellt werden.

#### **Erkenntnisse über allgemeines Twitter-Verhalten via Clickstream-Daten**

Bis dato hat noch keine Arbeit Logdaten oder Clickstream-Daten der Twitter-Website erhoben. Insofern stellt die methodische Herangehensweise sowie der erhaltene Datensatz ein Novum dar. Mithilfe des Datensatzes sind wesentliche Einblicke in allgemeines Twitter-Verhalten möglich, die mithilfe von qualitativen Studien, Experimenten oder API-Studien nicht zu erlangen sind. Dies betrifft beispielsweise Aussagen über: Länge und Häufigkeit von Twitter-Sitzungen, der Anzahl an Ereignissen pro Sitzung und den Besuch von verschiedenen Unterseiten. Diese Informationen können Aufschluss über Intentionen und Bedürfnisse von Nutzern bei der Interaktion mit Twitter geben.

#### **Nutzerbezogener Lebenszyklus von Tweets**

Die Arbeit zeigt, dass Tweets auch noch zwei Monate nach dem initialen Kontakt für einen Nutzer von Relevanz sein können und der nutzerbezogene Lebenszyklus damit länger ist, als man vermuten würde. Kontraintuitiv scheint dabei die Tatsache zu sein, dass, wenn Tweets ein initiales Relevanzsignal erhalten — durch ein Favorisieren oder einen RT — dies den Lebenszyklus verkürzt und nicht verlängert. Nutzer kehren also häufiger zu Tweets zurück, die sie beim ersten Aufeinandertreffen nicht als relevant erachtet haben. Zudem kann festgestellt werden, dass sehr viele ältere Tweets betrachtet werden. Trotz des Echtzeitfokus, den Twitter ohne Zweifel inne hat, zeigen diese beiden Erkenntnisse der Arbeit, dass Twitter auch Archivcharakter besitzt.

#### **Identifikation von Wiederfindensstrategien auf Twitter**

Obwohl Erkenntnisse aus Literatur und Ergebnisse der eigenen Arbeit die Annahmen unterstützen, dass Nutzer Tweets favorisieren oder weiterleiten, um diese später via dieser Speicher wiederzufin-

den, muss in der Logstudie festgestellt werden, dass diese Speicherorte nur sehr selten zum Wiederfinden genutzt werden. In der Diskussion in Abschnitt 7 wurden hier verschiedene Thesen aus der PIM-Forschung angeführt, um diese Beobachtungen der Logstudie zu interpretieren. Konkret konnten sowohl in der Umfrage als auch der Logstudie zwei bzw. drei Wiederfindensstrategien ausgemacht werden. Liegt das initiale Aufeinandertreffen mit dem gesuchten Tweet nur wenige Stunden zurück, so versuchen Nutzer via Scrolling diesen Tweet in der *Timeline* zu lokalisieren. Liegt das initiale Auffinden des Tweets länger zurück, so versuchen Nutzer den Tweet auf der Profilseite des Tweet-Autors zu lokalisieren. Hierfür verwenden sie zwei Orienteering-Strategien nämlich suchbasiertes *orienteering*, bei dem Profilseiten via Suche aufgesucht werden und navigationsbasiertes *orienteering*, bei dem Nutzer Links verwenden, um zu den Profilseiten zu gelangen. Reine Schlüsselwort-Suche, also Teleporting-Verhalten, kann nur in wenigen Fällen als Wiederfindensstrategie beobachtet werden.

#### **Entwicklung von Modellen zur Vorhersage von Wiederfindenszielen**

Im Kontext des nutzerbezogenen Lebenszyklus wurde bereits angesprochen, dass Nutzer nur bedingt dazu in der Lage sind, zukünftige Informationsbedürfnisse vorherzusagen. Um dem Nutzer diese fehleranfällige und schwierige Aufgabe abzunehmen, wurden mithilfe von maschinellen Lernverfahren Modelle entwickelt, die es ermöglichen, Tweets vorherzusagen, die potentiell wiedergefunden werden. Bereits mit Informationen über den Tweetinhalt und wie mit dem Tweet interagiert wurde, können bei drei von vier Tweets richtige Vorhersagen getroffen werden.

#### **Identifikation von starkem Personenfokus bei Twitterverhalten**

An mehreren Punkten der Arbeit wird deutlich: Personen und die Beziehung zwischen Tweetautor und Rezipienten sind von großer Bedeutung sowohl für Wiederfindensverhalten als auch für Twitterverhalten im Allgemeinen.

Es lässt sich erkennen, dass das Sequential-Mixed-Methods-Design, also die Ergänzung der qualitativen bzw. subjektiven Umfragedaten mit quantitativen bzw. objektiven Daten der Logstudie sich als sehr fruchtbar erwiesen hat, um Wiederfindensverhalten auf Twitter zu dokumentieren, beschreiben und interpretieren. Im Folgenden soll das Forschungsprojekt jedoch nochmals aus einer kritischen Perspektive betrachtet werden.

## **9.2. Kritische Anmerkungen zur Arbeit**

Kritische Anmerkungen können in Bezug auf die verwendete Methodik sowie auf Social-Media-Plattformen als Untersuchungsgegenstände gemacht werden.

### **9.2.1. In Bezug auf die verwendeten Methoden**

Kritik in Bezug auf die verwendeten Methoden ist gegenüber der Logstudie angebracht. Diese Form der Datenerhebung war beschränkt auf Interaktionen mit der Twitter-Website, die über die *Chrome*-Erweiterung erhoben wurden. Somit wurden keine Interaktionen von anderen Endgeräten oder Clients wie *TweetDeck* analysiert. Betrachtet man die Nutzungsdaten von Twitter, wird deutlich, dass etwa 80% den Service vor allem über mobile Endgeräte nutzen. Man kann folglich argumentieren, dass die Nutzung des Service via Website nicht das gewöhnliche Interaktionsszenario darstellt und sich Nutzer bei der Verwendung der mobilen Endgeräte sicherlich anders verhalten (in Bezug auf Häufigkeit, Länge sowie Art der Nutzung) als bei der Verwendung der Website. In der Arbeit wurde jedoch festgestellt, dass es viele Nutzer gibt, die — je nach Situation — Twitter sowohl über das mobile Endgerät als auch über die Website nutzen. Die Erhebung der Logdaten

erfolgte aber lediglich auf der Basis der Twitter-Website, so dass keine Cross-device-Interaktionen aufgezeichnet wurden. Diese stellen aber vermutlich klassisches Nutzerverhalten dar, wodurch sich für die Datenerhebung eine Lücke ergibt. Ein mögliches Szenario wäre folgendes: Nutzer lesen Tweets auf ihren mobilen Geräten und kehren für eine detaillierte Auseinandersetzung mit als relevant erachteten Tweets später über die Website zu diesen Tweets zurück. Viele Wiederfindensaktionen konnten somit nicht aufgezeichnet werden, da nur Interaktionen mit der Website geloggt wurden. Jedoch kann dieser Umstand auch positiv interpretiert werden: Obwohl von den Probanden keine Interaktionen mit Tweets über mehrere Endgeräte hinweg registriert wurden, konnten zahlreiche Wiederfindensaktionen in den Logs identifiziert werden. Dieser Umstand und die Tatsache, dass für die Analyse ein sehr konservativer Proxy zur Identifikation von Wiederfinden verwendet wurde, lässt vermuten, dass diese Arbeit die Häufigkeit und Problematik von Wiederfinden noch unterschätzt und Nutzer weitaus häufiger Tweets wiederfinden, als dies in dieser Arbeit gezeigt wird.

Ein weiterer Kritikpunkt bietet die Anzahl und der Umfang, mit dem Probanden an der Logstudie teilgenommen haben. Obwohl 44 Probanden die *Chrome*-Browsererweiterung heruntergeladen und installiert haben, ist ein gewisser *bias* einiger weniger Probanden, die für einen Großteil der Interaktionen in den Daten verantwortlich sind, nicht von der Hand zu weisen (vgl. Abschnitt 6.4.2). Eine größere und durch diverses Verhalten charakterisierte Stichprobe hätte durchaus zu weiteren Erkenntnissen in Bezug auf den nutzerbezogenen Lebenszyklus von Tweets sowie verwendeten Wiederfindensstrategien führen können.

### 9.2.2. In Bezug auf den Untersuchungsgegenstand

Bereits in Abschnitt 2.1.1 wurde darauf hingewiesen, dass die Untersuchung von Social-Media-Plattformen mit Problemen behaftet ist. Twitter ist ein lebender Organismus, der einem ständigen Wandel und Veränderungen unterliegt. Neben den Überarbeitungen, die sowohl am User-Interface als auch an der Funktionalität der Plattform in regelmäßigen Abständen durchgeführt werden, ist die Nutzungserfahrung einzelner Nutzer vermutlich auch dadurch beeinflusst, wenn diese Teil von A/B-Tests werden und z. B. deren *Timeline* aufgrund von Experimenten algorithmischen Veränderungen unterliegt. Auch andere Kontextfaktoren können einen Einfluss darauf nehmen, wie Twitter genutzt werden kann. Beispielsweise können Funktionalität je nach Land oder auch mobilem Betriebssystem variieren. Wie bereits in der Einleitung erwähnt wurden glücklicherweise keine für PIM-Verhalten relevanten Veränderungen an Twitter während der Studienzeit vorgenommen. Es kann jedoch nicht ausgeschlossen werden, dass einzelne Nutzer eine veränderte Twitter-Nutzungserfahrung dadurch erlebt haben, als sie Teil von A/B-Tests wurden.

Eine weitere Frage, die sich in diesem Zusammenhang aufdrängt, ist die Frage nach der Relevanz älterer Studien, die vor bestimmten Veränderungen von Twitter-Funktionalitäten durchgeführt wurden. Sind beispielsweise die Ergebnisse zur Untersuchung der RT-Funktion aus der 1. Phase noch von Bedeutung bzw. besitzen deren Erkenntnisse noch Gültigkeit für die aktuelle Nutzungssituation? Das Problem, dass Twitter eine lebende sich ständig weiterentwickelnde Plattform ist und die Tatsache, dass sich hieraus vermeintlich Probleme in Bezug auf die Aussagekraft von Studien ergeben, bleibt in der Wissenschaft nicht unberücksichtigt. So argumentieren Gorrell und Bontcheva: „We also need to acknowledge that Twitter usage patterns have been shown to be evolving rapidly and therefore previous literature does not necessarily apply to a current Twitter snapshot“ (Gorrell & Bontcheva, 2016, S.19).

Dieser These sollte sicherlich mit Studien nachgegangen werden. Ein Anknüpfungspunkt wird auch durch diese Arbeit geleistet, nämlich durch die Analyse der Favoriten-Funktion und den Beweggründen für das Favorisieren (vgl. Abschnitt 5.6). Im November 2015 wurde die Favoriten-

Funktion in die Like-Funktion umbenannt. Folglich wäre eine identische Studie in Bezug auf die Like-Funktionalität denkbar, um zu überprüfen, ob die Wahl eines anderen Icons und eine andere Bezeichnung das mentale Modell und die Nutzungsweise dieser Funktionalität geändert hat.

## 9.3. Anknüpfungspunkte für zukünftige Forschung

Abschnitt 9.3 stellt mögliche Anknüpfungspunkte für zukünftige Forschungsvorhaben vor. Forschungsarbeiten sind auf der Grundlage von drei Perspektiven möglich:

1. Aufbauend auf den bestehenden Logdaten unter Veränderungen von Wiederfindensdefinitionen bzw. Platzhaltermaßen (Proxies)
2. Aufbauend auf der *Chrome*-Browsererweiterung für die Evaluation der in Kapitel 8 vorgestellten Designvorschläge zur Unterstützung des Wiederfindensverhaltens
3. Aufbauend auf neuen Studien, die an offenen Fragen der Promotionsarbeit andocken und Forschungsdesiderate bilden

Da der zweiten Perspektive in Kapitel 8 bereits ausreichend Raum eingeräumt wurde, konzentrieren sich die folgenden Ausführungen auf die beiden anderen Perspektiven.

### 9.3.1. In Bezug auf existierende Daten

Bereits in der Einleitung der Arbeit wurde eine Differenzierung zwischen tatsächlichem Wiederfinden und verwandten Handlungen wie *re-visit* und *re-discovery* entlang der Dimension Zeit — also verstrichene Zeit zwischen initialem und erneutem Kontakt mit einem Informationsobjekt — vorgenommen. Neben Wiederfinden können in zukünftigen Studien auch Verhaltensweisen wie *re-visits* und *re-discovery* mithilfe der Logs untersucht werden. Im Fokus der Arbeit stand jedoch das Wiederfindensverhalten. In der Twitter-Logstudie wurde hierfür ein Platzhaltermaß verwendet, das Re-finding in den Logs mit möglichst hoher Genauigkeit identifizieren sollte. Es besteht ebenfalls die Möglichkeit, Re-finding mit etwas liberaleren Platzhaltermaßen wie in Abschnitt 6.4.4.2 beschrieben — z. B. mit *Klick* — *Hover* über Sessions hinweg — zu analysieren, um weitere Vergleiche zu den hier präsentierten Resultaten durchzuführen. So könnte z. B. analysiert werden, wie sich der zeitliche Rahmen des nutzerbezogenen Lebenszyklus sowie Wiederfindensstrategien verändern, wenn Wiederfinden via eines liberaleren Proxies definiert wird. Die Platzhaltermaße können ebenfalls für die beiden anderen Verhaltensweisen instrumentalisiert werden. Beispielsweise könnte eine erneute Interaktion wie *Hover* — *Hover* innerhalb einer Session zur Analyse von erneuten Besuchen genutzt werden. Für *re-discovery* könnte das 3. Quantil (21 Tage) des nutzerbezogenen Lebenszyklus/Wiederfindensperiode als zeitliche Dimension zu einem Proxy hinzugefügt werden, unter der Annahme, dass sich Nutzer nach dieser Zeitperiode nicht mehr tatsächlich an den Tweet erinnern, so dass es sich stattdessen mehr um ein erneutes Entdecken handelt als um ein bewusstes Wiederfinden.

Eine weitere Perspektive betrifft das Wiederfinden von eigenen Tweets. In der Einleitung wurde der Fokus der Arbeit auf Tweets aus der *Timeline* gelegt. Eigene gesendete Nachrichten wurden zwar nicht gänzlich vernachlässigt (siehe z. B. die Analyse der Wiederfindensgründe aus Abschnitt 5.8), jedoch wurde vor allem bei der quantitativen Analyse der Logdaten die Analyse des Wiederfindens von Tweets aus der *Timeline* in den Mittelpunkt gestellt. Somit eröffnet sich eine weitere Perspektive für zukünftige Studien, die darin besteht, die Analysen um das Wiederfinden von eigenen gesendeten Tweets zu erweitern und die Logdaten dahingehend auszuwerten. Eine weitere interessante Dimension, welche sich mit der Perspektive der eigenen gesendeten Tweets überschneidet,

besteht in der Analyse des *undo behaviour*, das das Löschen von eigenen gesendeten Nachrichten (siehe die Dimension der *regret deletion* aus Abschnitt 3.2.2), das Entfavorisieren eines Tweets aus der Favoritenliste oder das Zurücknehmen eines RTs untersucht. Bei allen drei Verhaltensweisen kann man davon ausgehen, dass ein Wiederfinden vorliegt. Diese Perspektive wurde in den vorliegenden Analysen der Logs nicht berücksichtigt, würde aber die Zahl der Wiederfindensfälle vergrößern.

### 9.3.2. In Bezug auf neue Studien

Neben weiteren Studien, die sich auf der Basis der bereits gesammelten Daten ergeben, können in einem nächsten Schritt weitere Studien durchgeführt werden.

#### Meta-Review zu Twitter-Forschungsarbeiten

Die Darstellung der relevanten Literatur zur Twitter-Forschung in Abschnitt 3 legt offen, dass noch kein ganzheitliches Literatur-Review bestehender Forschungsarbeiten zu Twitter durchgeführt wurde. Existierende Arbeiten von Williams, Terras und Warwick, Rogers oder Zimmer und Proferes gründen ihre Meta-Studien zur Twitter-Forschung oft nur auf einer geringen Zahl an Veröffentlichungen, besitzen eine thematisch zu eingeschränkte Perspektive und nehmen vorwiegend qualitative Inhaltsanalysen dieser Arbeiten vor (R. Rogers, 2013; Williams et al., 2013b; Zimmer & Proferes, 2014). Hier besteht sicherlich Raum für eine ganzheitliche Analyse der inhaltlichen Bandbreite mit der Twitter untersucht wird. Dies kann auf der Basis von quantitativen Verfahren, wie etwa einem Topic-Modelling-Ansatz angewendet auf Titel und Abstracts — welche aus mehreren verschiedenen Datenbanken wie Scopus, Web of Science, PubMed etc. zu einem Korpus aus mehreren Tausend Arbeiten zusammengetragen werden — realisiert werden.

#### Analyse von Re-finding via großangelegter Twitter-API-Studien

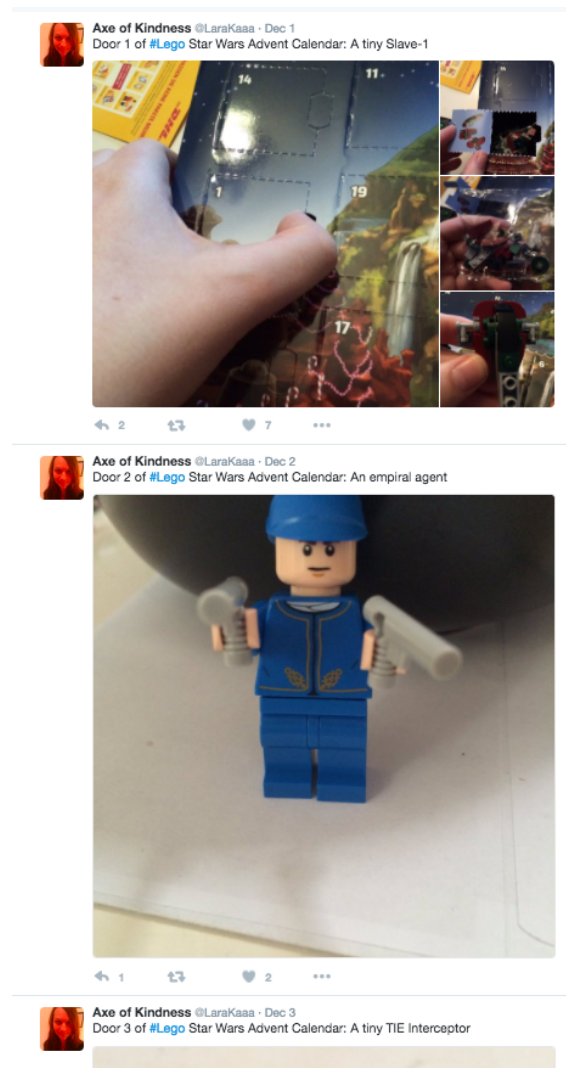
In den kritischen Anmerkungen zur Arbeit wurde bereits darauf hingewiesen, dass Twitter ein lebender Organismus ist, dessen Funktionalitäten einem ständigen Wandel unterliegen. Zwei Veränderungen an Twitter, welche 2016 durchgeführt wurden, erlauben es nun, Wiederfindensverhalten auch auf API-Basis zu untersuchen. Diese Veränderungen sind:

- Die neue @Reply-Funktionalität, die es ermöglicht, auf eigene Tweets zu antworten und dies sogar, ohne die textuelle @Mention — welche bis zur Veränderung der Reply-Funktion noch nötig war, um den Empfänger eines Tweets zu kennzeichnen — zu verwenden (Sherman, 2016).
- Die Veränderung der RT-Funktionalität, welche nun das Kommentieren von weitergeleiteten Tweets erlaubt (Shu, 2015).

Beide Funktionalitäten können dazu genutzt werden, um das 140-Zeichen-Limit Twitters indirekt zu umgehen, und thematisch zusammenhängende Threads aus einzelnen Tweets zu erstellen. Abbildung 9.1 zeigt exemplarisch wie das Vorgehen der *self replies* von einer Nutzerin dazu verwendet wurde, um sämtliche Inhalte eines Lego-Adventskalenders in einen zusammenhängen Post auf Twitter zu erstellen. Für jede der 24 @Replies musste ein Wiederfinden stattfinden.

Vordergründig stehen auch in diesem Kontext eigene gesendete Tweets im Fokus, da bei diesen die Annahme zutrifft, dass Nutzer diese wiederfinden müssen, um auf diese antworten oder mit Kommentar weiterleiten zu können. Eine Möglichkeit, um an eine große Menge an wiedergefundenen Tweets zu gelangen, besteht darin, die Twitter-Streaming-API zu überwachen und Tweets zu speichern, bei denen Autor und Antwortender bzw. Weiterleitender, den selben @Screenname besitzen. Dieser Tweet-Korpus kann dann als Basis für inhaltliche und zeitliche Analysen in Bezug auf Wiederfinden dienen.





**Abbildung 9.1.:** Visualisierung der Instrumentalisierung der neuen Reply-Funktion zur Erstellung eines Threads an zusammenhängenden Tweets

(Screenshot von: <https://twitter.com/larakaaa>)

### Cross-Plattform-Studie zu Wiederfindensverhalten

In Abschnitt 3.2.3 wurde im Rahmen von Studien zu Wiederfindensverhalten auf Social-Media-Plattformen die Arbeit von Oh et al. vorgestellt, welche das Phänomen des *re-tagging* von Instagram-Posts im Kontext von Fotowettbewerben wie dem *Weekend Hashtag Project* (#WHP) analysieren. Sie stellen fest, dass 35% aller Fotos ältere, wiedergefunden Fotos sind, die für die Teilnahme an den wöchentlich stattfindenden Wettbewerben mit einem neuen Tag versehen werden. Eine interessante Frage, die sich hierbei ergibt, ist die Frage, ob Nutzer, die dieses Verhalten auf Instagram an den Tag legen, ebenfalls dazu tendieren, sich auf Twitter ähnlich zu verhalten, indem sie die neue Kommentarfunktion des RT-Feature dazu nutzen, um ältere, eigene Tweets mit neuem Kommentar in den Umlauf zu bringen. Dies könnte zunächst generell, also ohne bestimmten inhaltlichen Bezug erfolgen, könnte aber auch im Kontext von Wettbewerben/Projekten wie dem #WHP durchgeführt werden, da vergleichbare Projekte auch auf Twitter existieren. Ziel hierbei ist es, eine Stichprobe

an Probanden zu finden, die sowohl über einen Instagram-Account als auch über einen Twitter-Account verfügen. Anschließend wird deren *re-tagging* bzw. *re-tweeting* Verhalten über einen längeren Zeitraum via API-Daten der beiden Plattformen beobachtet. Als Kontrollgruppe könnten schließlich auch Nutzer dienen, die zwar über Instagram- und Twitter-Account verfügen, jedoch nicht an Projekten wie dem #WHP teilnehmen, um festzustellen, wie unterschiedlich diese Gruppen *re-tagging* bzw. RT mit Kommentar-Verhalten an den Tag legen und somit Aussagen über Wiederfindensverhalten zu treffen.

#### **Welche Faktoren tragen zur Salienz von Twitter-Nutzern bei?**

An mehreren Stellen der Arbeit wurde deutlich, dass der Tweet-Autor bzw. -Sender und der Rezipient in einem besonderen Verhältnis zueinander stehen. Eine These bzw. ein Erklärungsansatz der Arbeit besteht in der First-impression-Hypothese, also der Annahme, dass Nutzer Informationen im Informationsraum entlang bestimmter Charakteristika priorisieren, weshalb sie sich an diese auch gut erinnern können und der Tatsache, dass Wiederfinden von Tweets zu einem Großteil via Lokalisierung des Tweets in Nutzerprofilseiten gelöst wird, was ein Erinnern des Tweet-Senders voraussetzt. Fraglich ist jedoch, wie diese Priorisierung entlang einzelner Nutzer erfolgt, also die Fragen: Was trägt zur Salienz von Twitter-Nutzer bei? Welche Rolle spielen beispielsweise Tweet-Frequenz, Frequenz mit der Profilbilder oder Namen geändert werden? Welche Rolle spielt der persönliche Bezug zwischen Sender und Rezipient für die Salienz? Ist Reziprozität in der Follower-Followee-Beziehung Voraussetzung für Salienz? All dies sind Fragen, die in zukünftigen Studien untersucht werden könnten. Methodisch gesehen bieten sich — im Gegensatz zu den vorherigen Forschungsvorhaben, die alle stark auf großangelegte, quantitative Studien via API-Daten abzielen — viele Möglichkeiten an, die zur Analyse dieser Fragen genutzt werden könnten. So könnten diese Fragen auch mithilfe von Experimenten bearbeitet werden.

In der Summe kann man feststellen, dass die Promotionsarbeit viele Fragen in Bezug auf das Wiederfindensverhalten auf Twitter zu beantworten vermag. Gleichzeitig werden jedoch auch viele Anknüpfungspunkte für zukünftige Forschungsvorhaben geschaffen.

# Literaturverzeichnis

- Abbar, S., Mejova, Y. & Weber, I. (2015). You Tweet What You Eat: Studying Food Consumption Through Twitter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (S. 3197–3206). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2702123.2702153>
- Abel, F., Celik, I., Houben, G.-J. & Siehndel, P. (2011). Leveraging the Semantics of Tweets for Adaptive Faceted Search on Twitter. In L. Aroyo et al. (Hrsg.), *The Semantic Web – ISWC 2011: 10th International Semantic Web Conference, Bonn, Germany, October 23-27, 2011, Proceedings, Part I* (S. 1–17). Berlin: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-3-642-25073-6\\_1](http://dx.doi.org/10.1007/978-3-642-25073-6_1)
- Abela, C. & Staff, C. (2016). Behaviour Mining for Automatic Task-Keeping and Visualisations for Task-Refinding. In *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval* (S. 23–32). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2854946.2854966>
- Abrams, D., Baecker, R. & Chignell, M. (1998). Information Archiving with Bookmarks: Personal Web Space Construction and Organization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 41–48). New York, NY, USA: ACM Press/Addison-Wesley Publishing Co. Zugriff auf <http://dx.doi.org/10.1145/274644.274651>
- Adar, E., Teevan, J. & Dumais, S. T. (2008). Large Scale Analysis of Web Revisitation Patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 1197–1206). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1357054.1357241>
- Agichtein, E., Brill, E. & Dumais, S. (2006). Improving Web Search Ranking by Incorporating User Behavior Information. In *Proceedings of SIGIR'06* (S. 19–26). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1148170.1148177>
- Agosti, M., Crivellari, F. & Di Nunzio, G. M. (2011). Web log analysis: a review of a decade of studies about information acquisition, inspection and interpretation of user interaction. *Data Mining and Knowledge Discovery*, 24 (3), 663–696. Zugriff auf <http://dx.doi.org/10.1007/s10618-011-0228-8>
- Agosto, D. E. & Hughes-Hassell, S. (2005). People, addresss, and questions: An investigation of the everyday life information-seeking behaviors of urban young adults. *Library & Information Science Research*, 27 (2), 141 – 163. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0740818805000046>
- Almuhimedi, H., Wilson, S., Liu, B., Sadeh, N. & Acquisti, A. (2013). Tweets Are Forever: A Large-scale Quantitative Analysis of Deleted Tweets. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (S. 897–908). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2441776.2441878>

- Alonso, O., Marshall, C. C. & Najork, M. (2013). Are Some Tweets More Interesting Than Others? #HardQuestion. In *Proceedings of HCIR'13* (S. 2:1–2:10). ACM. Zugriff auf <http://doi.acm.org/10.1145/2528394.2528396>
- André, P., Bernstein, M. & Luther, K. (2012). Who Gives a Tweet?: Evaluating Microblog Content Value. In *Proceedings of CSCW'12* (S. 471–474). ACM. Zugriff auf <http://doi.acm.org/10.1145/2145204.2145277>
- Antoun, C., Zhang, C., Conrad, F. G. & Schober, M. F. (2016). Comparisons of online recruitment strategies for convenience samples. *Field Methods*, 28 (3), 231-246. Zugriff auf <http://dx.doi.org/10.1177/1525822X15603149>
- Arguello, J. (2014). Predicting Search Task Difficulty. In M. de Rijke et al. (Hrsg.), *Proceedings of ECIR'14* (S. 88–99). Cham: Springer International Publishing. Zugriff auf [http://dx.doi.org/10.1007/978-3-319-06028-6\\_8](http://dx.doi.org/10.1007/978-3-319-06028-6_8)
- Bagdouri, M. & Oard, D. W. (2015). On Predicting Deletions of Microblog Posts. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (S. 1707–1710). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2806416.2806600>
- Baltar, F. & Brunet, I. (2012). Social research 2.0: virtual snowball sampling method using Facebook. *Internet Research*, 22 (1), 57-74. Zugriff auf <http://dx.doi.org/10.1108/10662241211199960>
- Barbaro, M. & Zeller, T. J. (2006). *A Face Is Exposed for AOL Searcher No. 4417749*. Zugriff auf <http://query.nytimes.com/gst/abstract.html?res=9E0CE3DD1F3FF93AA3575BC0A9609C8B63>
- Barnes, S. J. & Böhringer, M. (2011). Modeling Use Continuance Behavior in Microblogging Services: The Case of Twitter. *Journal of Computer Information Systems*, 51 (4), 1-10. Zugriff auf <http://www.tandfonline.com/doi/abs/10.1080/08874417.2011.116454960>
- Barreau, D., Capra, R., Dumais, S., Jones, W. & Pérez-Quinones, M. (2008). Introduction to Keeping, Refinding and Sharing Personal Information. *ACM Trans. Inf. Syst.*, 26 (4), 18:1–18:3. Zugriff auf <http://doi.acm.org/10.1145/1402256.1402257>
- Barreau, D. & Nardi, B. A. (1995). Finding and Reminding: File Organization from the Desktop. *SIGCHI Bull.*, 27 (3), 39–43. Zugriff auf <http://doi.acm.org/10.1145/221296.221307>
- Bates, M. J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science*, 50 (12), 1043–1050. Zugriff auf [http://dx.doi.org/10.1002/\(SICI\)1097-4571\(1999\)50:12<1043::AID-ASI1>3.0.CO;2-X](http://dx.doi.org/10.1002/(SICI)1097-4571(1999)50:12<1043::AID-ASI1>3.0.CO;2-X)
- Behrend, T. S., Sharek, D. J., Meade, A. W. & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior Research Methods*, 43 (3), 800–813. Zugriff auf <http://dx.doi.org/10.3758/s13428-011-0081-0>
- Bender, R. & Lange, S. (2007). Was ist ein Konfidenzintervall? *Deutsche Medizinische Wochenschrift*, 132, e17–e18. Zugriff auf <https://www.thieme-connect.de/products/ejournals/pdf/10.1055/s-2007-959031.pdf>
- Benevenuto, F., Rodrigues, T., Cha, M. & Almeida, V. (2009). Characterizing User Behavior in Online Social Networks. In *Proceedings of IMC'09* (S. 49–62). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1644893.1644900>
- Benevenuto, F., Rodrigues, T., Cha, M. & Almeida, V. (2012). Characterizing user navigation and interactions in online social networks. *Information Sciences*, 195, 1 - 24. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0020025511006372>

- Benn, Y., Bergman, O., Glazer, L., Arent, P., Wilkinson, I. D., Varley, R. & Whittaker, S. (2015). Navigating through digital folders uses the same brain structures as real world navigation. *Nature Scientific Reports*, 5, 14719 EP -. Zugriff auf <http://dx.doi.org/10.1038/srep14719>
- Berendsen, R., Tsagkias, M., Weerkamp, W. & de Rijke, M. (2013). Pseudo Test Collections for Training and Tuning Microblog Rankers. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 53–62). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2484028.2484063>
- Bergman, O., Beyth-Marom, R. & Nachmias, R. (2003). The user-subjective approach to personal information management systems. *Journal of the American Society for Information Science and Technology*, 54 (9), 872–878. Zugriff auf <http://dx.doi.org/10.1002/asi.10283>
- Bergman, O., Beyth-Marom, R. & Nachmias, R. (2008). The user-subjective approach to personal information management systems design: Evidence and implementations. *Journal of the American Society for Information Science and Technology*, 59 (2), 235–246. Zugriff auf <http://dx.doi.org/10.1002/asi.20738>
- Bergman, O., Beyth-Marom, R., Nachmias, R., Gradovitch, N. & Whittaker, S. (2008). Improved Search Engines and Navigation Preference in Personal Information Management. *ACM Trans. Inf. Syst.*, 26 (4), 20:1–20:24. Zugriff auf <http://doi.acm.org/10.1145/1402256.1402259>
- Bergman, O. & Whittaker, S. (2016). *The Science of Managing Our Digital Stuff*. Cambridge, London: The MIT Press.
- Bergman, O., Whittaker, S., Sanderson, M., Nachmias, R. & Ramamoorthy, A. (2010). The Effect of Folder Structure on Personal File Navigation. *Journal of the American Society for Information Science and Technology*, 61 (12), 2426–2441. Zugriff auf <http://dx.doi.org/10.1002/asi.v61:12>
- Bevans, G. E. (1913). *How Workingmen Spend Their Spare Time* (Dissertation, Columbia University, New York). Zugriff auf <https://archive.org/details/howworkingmenspe00bevarich>
- Boardman, R. & Sasse, M. A. (2004). „Stuff Goes into the Computer and Doesn’t Come out“: A Cross-tool Study of Personal Information Management. In *Proceedings of CHI’04* (S. 583–590). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/985692.985766>
- Bogers, T. & Björneborn, L. (2013). Micro-serendipity: Meaningful Coincidences in Everyday Life Shared on Twitter. In *Proc. of iConference’13* (S. 196–208). Zugriff auf <https://www.ideals.illinois.edu/bitstream/handle/2142/36052/175.pdf>
- Bolger, N. & Laurenceau, J.-P. (2013). *Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research*. New York, USA: Guilford Publications.
- Bollen, J., Mao, H. & Zeng, X. (2010). Twitter mood predicts the stock market. *CoRR*, abs/1010.3003. Zugriff auf <http://arxiv.org/abs/1010.3003>
- Booten, K. (2016). Hashtag Drift: Tracing the Evolving Uses of Political Hashtags Over Time. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (S. 2401–2405). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2858036.2858398>
- Borgman, C. L. (1986). Why Are Online Catalogs Hard to Use? Lessons Learned from Information Retrieval Studies. *Journal of the American Society for Information Science*, 37 (6), 387–400.
- Borgman, C. L. & Rice, R. E. (1982). Computer-Monitored Communication Data: Prospects and Problems for Communication Research. In *Information Interaction: Proceedings of the 45th ASIS Annual Meeting* (S. 37–40). White Plains, NY: Knowledge Industry Publications.

- Borondo, J., Morales, A., Benito, R. & Losada, J. (2014). Mapping the online communication patterns of political conversations. *Physica A: Statistical Mechanics and its Applications*, 414, 403 - 413. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0378437114006529>
- Bortz, J. & Doering, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Schuster, C. (2011). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer.
- boyd, D., Golder, S. & Lotan, G. (2010). Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In *Proceedings of HICSS'10* (S. 1–10). Washington, DC, USA: IEEE. Zugriff auf <http://dx.doi.org/10.1109/HICSS.2010.412>
- Brandstätter, E. (1999). Confidence Intervals as an Alternative to Significance Testing. *Methods of Psychological Research - Online*, 4 (2). Zugriff auf <http://www.dgps.de/fachgruppen/methoden/mpr-online/issue7/art2/brandstaetter.pdf>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45 (1), 5–32. Zugriff auf <http://dx.doi.org/10.1023/A:1010933404324>
- Breuer, F. (2009). *Reflexive Grounded Theory. Eine Einführung für die Forschungspraxis*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bruce, H. (2005). Personal, anticipated information need. *Information Research*, 10(3) (232). Zugriff auf <http://InformationR.net/ir/10-3/paper232.html>
- Bruce, H., Jones, W. & Dumais, S. (2004a). Information behaviour that keeps found things found. *Information Research*, 10 (1). Zugriff auf <http://www.informationr.net/ir/10-1/paper207.html>
- Bruce, H., Jones, W. & Dumais, S. (2004b). Keeping and re-finding information on the web: What do people do and what do they need? *Proceedings of the American Society for Information Science and Technology*, 41 (1), 129–137. Zugriff auf <http://dx.doi.org/10.1002/meet.1450410115>
- Bruns, A., Highfield, T. & Burgess, J. (2013). The Arab Spring and Social Media Audiences: English and Arabic Twitter Users and Their Networks. *American Behavioral Scientist*, 57 (7), 871-898. Zugriff auf <http://abs.sagepub.com/content/57/7/871.abstract>
- Buhrmester, M., Kwang, T. & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6 (1), 3-5. Zugriff auf <http://pps.sagepub.com/content/6/1/3.abstract>
- Burghardt, M., Karsten, H., Pflamminger, M. & Wolff, C. (2013). Twitter als interaktive Erweiterung des Mediums Fernsehen: Inhaltliche Analyse von Tatort-Tweets. In *Workshop Proceedings of the 25th Conference of the German Society for Computational Linguistics (GSCL 2013)*. Zugriff auf <http://epub.uni-regensburg.de/33807/>
- Busing, F. M. T. A., Weaver, B. & Dubois, S. (2016). 2 × 2 Tables: a note on Campbell's recommendation. *Statistics in Medicine*, 35 (8), 1354–1358. Zugriff auf <http://dx.doi.org/10.1002/sim.6808>
- Büttner, R. & Büttner, K. (2016). A Systematic Literature Review of Twitter Research from a Socio-Political Revolution Perspective. In *Proceedings of HICSS'16*. Washington, DC, USA: IEEE. Zugriff auf <http://ieeexplore.ieee.org/document/7427460/>

- Campbell, I. (2007). Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. *Statistics in Medicine*, 26 (19), 3661–3675. Zugriff auf <http://dx.doi.org/10.1002/sim.2832>
- Capra, R. & Pérez-Quñones, M. A. (2003). Re-Finding Found Things: An Exploratory Study of How Users Re-Find Information. *CoRR*, cs.HC/0310011. Zugriff auf <http://arxiv.org/abs/cs.HC/0310011>
- Capra, R. & Perez-Quinones, M. (2005). Using Web Search Engines to Find and Refind Information. *Computer*, 38 (10), 36–42.
- Capra, R., Pinney, M. & Perez-Quinones, M. (2005). *Refinding is Not Finding Again* (Bericht). Zugriff auf <http://eprints.cs.vt.edu/archive/00000715/>
- Carmel, D., Halawi, G., Lewin-Eytan, L., Maarek, Y. & Raviv, A. (2015). Rank by Time or by Relevance?: Revisiting Email Search. In *Proceedings of CIKM'15* (S. 283–292). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2806416.2806471>
- Carmody, T. (2011). *Can 'Serendipity' Be a Business Model? Consider Twitter*. Zugriff auf <http://www.wired.com/2011/11/serendipity-ads-twitter/all/>
- Cartright, M.-A., White, R. W. & Horvitz, E. (2011). Intentions and Attention in Exploratory Health Search. In *Proceedings of SIGIR'11* (S. 65–74). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2009916.2009929>
- Case, D. (2007). *Looking for Information* (2. Aufl.). London: Academic Press, Elsevier.
- Case, D. & Given, L. (2016). *Looking for Information* (4. Aufl.). Bingley: Emerald Group Publishing.
- Casler, K., Bickel, L. & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29 (6), 2156 - 2160. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S074756321300160X>
- Chang, Y., Tang, L., Inagaki, Y. & Liu, Y. (2014). What is Tumblr: A Statistical Overview and Comparison. *SIGKDD Explor. NewsL.*, 16 (1), 21–29. Zugriff auf <http://doi.acm.org/10.1145/2674026.2674030>
- Chau, D. H., Myers, B. & Faulring, A. (2008). What to Do when Search Fails: Finding Information by Association. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 999–1008). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1357054.1357208>
- Chen, A. T. (2015). Information use and illness representations: Understanding their connection in illness coping. *Journal of the Association for Information Science and Technology*, 66 (2), 340–353. Zugriff auf <http://dx.doi.org/10.1002/asi.23173>
- Chen, J., Nairn, R. & Chi, E. (2011). Speak Little and Well: Recommending Conversations in Online Social Streams. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 217–226). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1978942.1978974>
- Cherubini, M. & Oliver, N. (2009). A Refined Experience Sampling Method to Capture Mobile User Experience. *CoRR*, abs/0906.4125. Zugriff auf <http://arxiv.org/abs/0906.4125>
- Choi, B. C. & Pak, A. W. (2005). A Catalog of Biases in Questionnaires. *Preventing Chronic Disease*, 2 (1), A13. Zugriff auf <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1323316/>
- Choudhury, M. D., Gamon, M., Counts, S. & Horvitz, E. (2013). Predicting Depression via Social Media. In *Proceedings of ICWSM'13*. Washington: The AAAI Press. Zu-

- griff auf <https://www.microsoft.com/en-us/research/publication/predicting-depression-via-social-media/>
- Church, K., Cherubini, M. & Oliver, N. (2014). A Large-scale Study of Daily Information Needs Captured in Situ. *ACM Trans. Comput.-Hum. Interact.*, 21 (2), 10:1–10:46. Zugriff auf <http://doi.acm.org/10.1145/2552193>
- Cochrane, P. A. & Markey, K. (1983). Catalog Use Studies – Since the Introduction of Online Interactive Catalogs: Impact on Design for Subject Access. *Library & Information Science Research*, 5 (4), 337–363.
- Cockburn, A., Greenberg, S., Jones, S., McKenzie, B. & Moyle, M. (2003). Improving Web Page Revisitation: Analysis, Design and Evaluation. *IT&SOCIETY*, 1 (3), 159–183.
- Cockburn, A. & McKenzie, B. (2001). What do web users do? An empirical analysis of web use. *International Journal of Human-Computer Studies*, 54 (6), 903 - 922. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S1071581901904598>
- Cole, I. (1982). Human Aspects of Office Filing: Implications for the Electronic Office. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 26 (1), 59-63. Zugriff auf <http://pro.sagepub.com/content/26/1/59.abstract>
- Colleoni, E., Rozza, A. & Arvidsson, A. (2014). Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication*, 64 (2), 317–332. Zugriff auf <http://dx.doi.org/10.1111/jcom.12084>
- Collier, N., Son, N. T. & Nguyen, N. M. (2011). OMG U got flu? Analysis of shared health messages for bio-surveillance. *CoRR*, abs/1110.3089. Zugriff auf <http://arxiv.org/abs/1110.3089>
- Consolvo, S. & Walker, M. (2003). Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing*, 2 (2), 24–31. Zugriff auf <http://dx.doi.org/10.1109/MPRV.2003.1203750>
- Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1 (1), 104. Zugriff auf <http://dx.doi.org/10.1007/BF03177550>
- Corbin, J. & Strauss, A. (2008). *Basics of qualitative Research* (3. Aufl.). Los Angeles: Sage.
- Cosijn, E. & Ingwersen, P. (2000). Dimensions of Relevance. *Inf. Process. Manage.*, 36 (4), 533–550. Zugriff auf [http://dx.doi.org/10.1016/S0306-4573\(99\)00072-2](http://dx.doi.org/10.1016/S0306-4573(99)00072-2)
- Counts, S. & Fisher, K. (2011). Taking It All In? Visual Attention in Microblog Consumption. In *Proceedings of ICWSM'11*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2769>
- Creswell, J. W. (2014). *Research Design. Qualitative, Quantitative & Mixed Methods Approaches* (4. Aufl.). Los Angeles: Sage.
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. & Hanson, W. (2003). Advanced Mixed Methods Research Designs. In C. Tashakkori A & Teddlie (Hrsg.), *Handbook of Mixed Methods in Social & Behavioral Research* (S. 209–240). Los Angeles: Sage.
- Csikszentmihályi, M. & Larson, R. (2014). The Experience Sampling Method. In *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (S. 21–34). Dordrecht: Springer Netherlands. Zugriff auf [http://dx.doi.org/10.1007/978-94-017-9088-8\\_3](http://dx.doi.org/10.1007/978-94-017-9088-8_3)
- Cui, A., Zhang, M., Liu, Y., Ma, S. & Zhang, K. (2012). Discover Breaking Events with Popular Hashtags in Twitter. In *Proceedings of the 21st ACM International Conference on Information*



- and Knowledge Management (S. 1794–1798). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2396761.2398519>
- Cunha, E., Magno, G., Almeida, V., Gonçalves, M. A. & Benevenuto, F. (2012). A Gender Based Study of Tagging Behavior in Twitter. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media* (S. 323–324). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2309996.2310055>
- Cunha, E., Magno, G., Comarela, G., Almeida, V., Gonçalves, M. A. & Benevenuto, F. (2011). Analyzing the Dynamic Evolution of Hashtags on Twitter: A Language-based Approach. In *Proceedings of the Workshop on Languages in Social Media* (S. 58–65). Stroudsburg, PA, USA: Association for Computational Linguistics. Zugriff auf <http://dl.acm.org/citation.cfm?id=2021109.2021117>
- Cutrell, E., Robbins, D., Dumais, S. & Sarin, R. (2006). Fast, Flexible Filtering with Phlat. In *Proceedings of CHI'06* (S. 261–270). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1124772.1124812>
- Dann, S. (2010). Twitter content classification. *First Monday*, 15 (12). Zugriff auf <http://journals.uic.edu/ojs/index.php/fm/article/view/2745>
- Dantonio, L., Makri, S. & Blandford, A. (2012). Coming across academic social media content serendipitously. *Proc. the American Society for Information Science and Technology*, 49 (1), 1–10. Zugriff auf <http://dx.doi.org/10.1002/meet.14504901002>
- Das, S. & Kramer, A. (2013). Self-Censorship on Facebook. In *Proceedings of ICWSM'13*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6093/6350>
- Dawson, R. & Bynghall, S. (2012). *Getting results from crowds*. San Francisco: Advanced Human Technologies.
- Deng, T., Zhao, L., Feng, L. & Xue, W. (2011). Information Re-finding by Context: A Brain Memory Inspired Approach. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (S. 1553–1558). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2063576.2063799>
- den Poel, D. V. & Buckinx, W. (2005). Predicting online-purchasing behaviour. *European Journal of Operational Research*, 166 (2), 557 - 575. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0377221704002875>
- Diekema, A. R. & Olsen, M. W. (2014). Teacher Personal information management (PIM) practices: Finding, keeping, and Re-Finding information. *Journal of the Association for Information Science and Technology*, 65 (11), 2261–2277. Zugriff auf <http://dx.doi.org/10.1002/asi.23117>
- DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B. & Muller, M. (2008). Motivations for social networking at work. In *Proceedings of the 2008 acm conference on computer supported cooperative work* (S. 711–720). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1460563.1460674>
- Dorsey, J. (2006). *twtrr sketch*. Zugriff auf <https://www.flickr.com/photos/jackdorsey/182613360/am21.12.2016>
- Duan, Y., Jiang, L., Qin, T., Zhou, M. & Shum, H.-Y. (2010). An Empirical Study on Learning to Rank of Tweets. In *Proceedings of the 23rd International Conference on Computational Linguistics* (S. 295–303). Stroudsburg, PA, USA: Association for Computational Linguistics. Zugriff auf <http://dl.acm.org/citation.cfm?id=1873781.1873815>

- Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R. & Robbins, D. C. (2003). Stuff I've Seen: A System for Personal Information Retrieval and Re-use. In *Proceedings of SIGIR'03* (S. 72–79). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/860435.860451>
- Dumais, S., Jeffries, R., Russell, D. M., Tang, D. & Teevan, J. (2014). Understanding User Behavior Through Log Data and Analysis. In J. S. Olson & W. A. Kellogg (Hrsg.), *Ways of Knowing in HCI* (S. 349–372). Berlin: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-1-4939-0378-8\\_14](http://dx.doi.org/10.1007/978-1-4939-0378-8_14)
- Efron, M. (2010). Hashtag Retrieval in a Microblogging Environment. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 787–788). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1835449.1835616>
- Efron, M. (2011). Information Search and Retrieval in Microblogs. *Journal of the American Society for Information Science and Technology*, 62 (6), 996–1008. Zugriff auf <http://dx.doi.org/10.1002/asi.21512>
- Efron, M., Lin, J., He, J. & de Vries, A. (2014). Temporal Feedback for Tweet Search with Non-parametric Density Estimation. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval* (S. 33–42). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2600428.2609575>
- Efron, M. & Winget, M. (2010). Questions are content: A taxonomy of questions in a microblogging environment. *Proceedings of the American Society for Information Science and Technology*, 47 (1), 1–10. Zugriff auf <http://dx.doi.org/10.1002/meet.14504701208>
- Ellis, D. (1989). A Behavioral Approach to Information Retrieval System Design. *J. Doc.*, 45 (3), 171–212. Zugriff auf <http://dx.doi.org/10.1108/eb026843>
- Ellis, D. (2011). The Emergence of Conceptual Modelling in Information Behaviour Research. In *New Directions in Information Behaviour* (S. 17–35). Bingley, UK: Emerald. Zugriff auf <http://www.emeraldinsight.com/doi/abs/10.1108/S1876-0562%282011%29002011a005>
- Elo, S. & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62 (1), 107–115. Zugriff auf <http://dx.doi.org/10.1111/j.1365-2648.2007.04569.x>
- Elsweiler, D., Baillie, M. & Ruthven, I. (2008). Exploring Memory in Email Refinding. *ACM Trans. Inf. Syst.*, 26 (4), 21:1–21:36. Zugriff auf <http://doi.acm.org/10.1145/1402256.1402260>
- Elsweiler, D., Baillie, M. & Ruthven, I. (2011). What Makes Re-finding Information Difficult? A Study of Email Re-finding. In *Proceedings of ECIR'13* (S. 568–579). Berlin: Springer. Zugriff auf <http://dl.acm.org/citation.cfm?id=1996889.1996963>
- Elsweiler, D. & Hammwöhner, R. (2015). Information Behavior – Ein zentrales Forschungsthema der Informationswissenschaft. *Information - Wissenschaft & Praxis*, 66 (1), 3–9.
- Elsweiler, D. & Harvey, M. (2015). Engaging and maintaining a sense of being informed: Understanding the tasks motivating twitter search. *Journal of the Association for Information Science and Technology*, 66 (2), 264–281. Zugriff auf <http://dx.doi.org/10.1002/asi.23182>
- Elsweiler, D., Harvey, M. & Hacker, M. (2011). Understanding Re-finding Behavior in Naturalistic Email Interaction Logs. In *Proceedings of SIGIR'11* (S. 35–44). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2009916.2009925>

- Elsweiler, D. & Ruthven, I. (2007). Towards Task-based Personal Information Management Evaluations. In *Proceedings of SIGIR'07* (S. 23–30). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1277741.1277748>
- Esposti, C. (2011). *Crowdsourcing Industry Landscape* Verfügbar unter: Zugriff auf [http://www.executiveaccelerators.com.au/\\_literature\\_129711/Crowd\\_Sourcing\\_Industry\\_Landscape\\_V09\\_Password](http://www.executiveaccelerators.com.au/_literature_129711/Crowd_Sourcing_Industry_Landscape_V09_Password)
- Fabritius, H. (1999). Triangulation as a multi-perspective strategy in a qualitative study of information seeking behaviour of journalists. In T. D. Wilson & D. K. Allen (Hrsg.), *Exploring the Contexts of Information Behaviour* (S. 406–419). London, UK: Taylor Graham Publishing. Zugriff auf <http://dl.acm.org/citation.cfm?id=315846.315874>
- Feng, W. & Wang, J. (2013). Retweet or Not?: Personalized Tweet Re-ranking. In *Proceedings of WSDM'13* (S. 577–586). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2433396.2433470>
- Fernández-Delgado, M., Cernadas, E., Barro, S. & Amorim, D. (2014). Do We Need Hundreds of Classifiers to Solve Real World Classification Problems? *Journal of Machine Learning Research*, 15 (1).
- Ferrara, E., Interdonato, R. & Tagarelli, A. (2014). Online Popularity and Topical Interests Through the Lens of Instagram. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media* (S. 24–34). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2631775.2631808>
- Fidel, R. (2008). Are we there yet?: Mixed methods research in library and information science. *Library & Information Science Research*, 30 (4), 265 - 272. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S074081880800073X>
- Field, A. & Miles, Z., J. & Field. (2014). *Discovering Statistics Using R*. Los Angeles: Sage.
- Fingas, J. (2016). *Twitter takes down a site that archives popular users' tweets*. Zugriff auf [https://www.engadget.com/2016/07/11/twitter-takes-down-verified-tweet-archive/?sr\\_source=Twitter](https://www.engadget.com/2016/07/11/twitter-takes-down-verified-tweet-archive/?sr_source=Twitter)
- Finneran, C. M. (2008). Factors that influence users to leave, acquire, and retain information items: A case study of college students' Personal Information Management. *Proceedings of the American Society for Information Science and Technology*, 45 (1), 1–5. Zugriff auf <http://dx.doi.org/10.1002/meet.2008.1450450365>
- Fisher, D., Brush, A. J., Gleave, E. & Smith, M. A. (2006). Revisiting Whittaker & Sidner's "Email Overload" Ten Years Later. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (S. 309–312). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1180875.1180922>
- Fisher, K., Erdelez, S. & McKechnie, L. (2005). *Theories of Information Behavior*. Medford: Information Today, Inc.
- Fitchett, S., Cockburn, A. & Gutwin, C. (2014). Finder Highlights: Field Evaluation and Design of an Augmented File Browser. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 3685–3694). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557014>
- Fitton, L., Gruen, M. & Poston, L. (2010). *Twitter For Dummies* (2. Aufl.). For Dummies.
- Fowler, F. J. J. & Cosenza, C. (2008). Writing Effective Questions. In E. D. de Leeuw, J. Hox & D. Dillman (Hrsg.), *International Handbook on Survey Research* (S. 136–159). London: Routledge.

- Fox, B. I. & Varadarajan, R. (2011). Use of Twitter to Encourage Interaction in a Multi-campus Pharmacy Management Course. *American Journal of Pharmaceutical Education* (5), 88. Zugriff auf <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3142990/>
- Froehlich, J., Chen, M. Y., Consolvo, S., Harrison, B. & Landay, J. A. (2007). MyExperience: A System for in Situ Tracing and Capturing of User Feedback on Mobile Phones. In *Proceedings of mobisys'07* (S. 57–70). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1247660.1247670>
- Fuchs, C. & Groh, G. (2015). Appropriateness of Search Engines, Social Networks, and Directly Approaching Friends to Satisfy Information Needs. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (S. 1248–1253). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2808797.2808836>
- Gadiraju, U., Kawase, R., Dietze, S. & Demartini, G. (2015). Understanding Malicious Behavior in Crowdsourcing Platforms: The Case of Online Surveys. In *Proceedings of CHI'15* (S. 1631–1640). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2702123.2702443>
- Gaffney, D. (2010). #iranElection: quantifying online activism. In *Proceedings of the WebSci10: Extending the Frontiers of Society On-Line*. Raleigh, NC. Zugriff auf [http://journal.webscience.org/295/2/websci10\\_submission\\_6.pdf](http://journal.webscience.org/295/2/websci10_submission_6.pdf)
- Galuba, W., Aberer, K., Chakraborty, D., Despotovic, Z. & Kellerer, W. (2010). Outtweeting the Twitterers - Predicting Information Cascades in Microblogs. In *Proceedings of WOSN'10* (S. 3–3). USENIX Association. Zugriff auf <http://dl.acm.org/citation.cfm?id=1863190.1863193>
- Garimella, K., Weber, I. & Cin, S. D. (2014). From „i love you babe“ to „leave me alone“- romantic relationship breakups on twitter. *CoRR*, abs/1409.5980. Zugriff auf <http://arxiv.org/abs/1409.5980>
- Geißler, C. (2010). Was sind ... Social Media? *Harvard Business Manager*, 9. Zugriff auf <http://www.webcitation.org/6EcpIOpal>
- Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Gläser, J. & Laudel, G. (2009). *Experteninterviews und qualitative Inhaltsanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gleason, B. (2013). #Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter. *American Behavioral Scientist*. Zugriff auf <http://abs.sagepub.com/content/early/2013/03/15/0002764213479372.abstract>
- Golbeck, J., Grimes, J. M. & Rogers, A. (2010). Twitter use by the U.S. Congress. *Journal of the American Society for Information Science and Technology*, 61 (8), 1612–1621. Zugriff auf <http://dx.doi.org/10.1002/asi.21344>
- Gomez-Rodriguez, M., Gummadi, K. & Schölkopf, B. (2014). Quantifying Information Overload in Social Media and its Impact on Social Contagions. In *Proceedings of ICWSM '14*. Washington: The AAAI Press. Zugriff auf <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8108/8117>
- Google. (2016). *Hosting*. Zugriff auf <https://developer.chrome.com/extensions/hosting>

- Gorrell, G. & Bontcheva, K. (2016). Classifying Twitter favorites: Like, bookmark, or Thanks? *Journal of the Association for Information Science and Technology*, 67 (1), 17–25. Zugriff auf <http://dx.doi.org/10.1002/asi.23352>
- Gottfried, J. & Shearer, E. (2016). *News Use Across Social Media Platforms 2016*. Zugriff auf <http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>
- Grbovic, M., Halawi, G., Karnin, Z. & Maarek, Y. (2014). How Many Folders Do You Really Need? Classifying Email into a Handful of Categories. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (S. 869–878). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2661829.2662018>
- Greenwood, S., Perrin, A. & Duggan, M. (2016). *Social Media Update 2016*. Zugriff auf <http://www.pewinternet.org/2016/11/11/social-media-update-2016/>
- Greifeneder, E. (2014). Trends in information behaviour research. *Inf. Res.*, 19 (4). Zugriff auf <http://www.informationr.net/ir/19-4/isic/isic13.html>
- Grevet, C., Choi, D., Kumar, D. & Gilbert, E. (2014). Overload is Overloaded: Email in the Age of Gmail. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 793–802). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557013>
- Griffiths, J. R., Hartley, R. J. & Willson, J. P. (2002). An improved method of studying user-system interaction by combining transaction log analysis and protocol analysis. *Inf. Res.*, 7 (4). Zugriff auf <http://informationr.net/ir/7-4/paper139.html>
- Grouven, U., Bender, R., Ziegler, A. & Lange, S. (2007). Der Kappa-Koeffizient. *Deutsche medizinische Wochenschrift*, 132 (S 01), e65-e68. Zugriff auf <https://www.thieme-connect.com/products/ejournals/pdf/10.1055/s-2007-959046.pdf>
- Gruzd, A. & Roy, J. (2014). Investigating Political Polarization on Twitter: A Canadian Perspective. *Policy & Internet*, 6 (1), 28–45. Zugriff auf <http://dx.doi.org/10.1002/1944-2866.POI354>
- Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L. & Albatal, R. (2016). NTCIR Lifelog: The First Test Collection for Lifelog Research. In *Proceedings of SIGIR'16* (S. 705–708). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2911451.2914680>
- Hair, J., Black, W. & Babin, B. (2010). *Multivariate Data Analysis: A Global Perspective*. Upper Saddle River, NJ, USA: Pearson Education.
- Hall, C. & Zarro, M. (2012). Social curation on the website Pinterest.com. *Proceedings of the American Society for Information Science and Technology*, 49 (1), 1–9. Zugriff auf <http://dx.doi.org/10.1002/meet.14504901189>
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in quantitative methods for psychology*, 8 (1), 23–34.
- Han, J., Choi, D., Chun, B.-G., Kwon, T., Kim, H.-c. & Choi, Y. (2014). Collecting, Organizing, and Sharing Pins in Pinterest: Interest-driven or Social-driven? *SIGMETRICS Perform. Eval. Rev.*, 42 (1), 15–27. Zugriff auf <http://doi.acm.org/10.1145/2637364.2591996>
- Hansen, F. C. (1980). The Process of Searching Online Bibliographic Databases: A Review of Research. *Library Research*, 2 (2), 107-127.
- Harrell, F. E. (2006). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Berlin: Springer.

- Harris, J. K. & Burke, R. C. (2005). Do you see what I see? An application of inter-coder reliability in qualitative analysis. In *American Public Health Association. 133rd Annual Meeting & Exposition*. Washington.
- Harvey, M. & Elswiler, D. (2012). Exploring Query Patterns in Email Search. In R. Baeza-Yates et al. (Hrsg.), *Proceedings of ECIR'12* (S. 25–36). Berlin: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-3-642-28997-2\\_3](http://dx.doi.org/10.1007/978-3-642-28997-2_3)
- Hasanain, M., Bagdouri, M., Elsayed, T. & Oard, D. (2016). What Questions Do Journalists Ask on Twitter? In *Proceedings of ICWSM'16*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13221>
- Hasanain, M., Elsayed, T. & Magdy, W. (2014). Identification of Answer-Seeking Questions in Arabic Microblogs. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (S. 1839–1842). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2661829.2661959>
- Hassan, A., White, R. W., Dumais, S. T. & Wang, Y.-M. (2014). Struggling or Exploring?: Disambiguating Long Search Sessions. In *Proceedings of WSDM'14* (S. 53–62). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556195.2556221>
- Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction* (2. Aufl.). Berlin: Springer.
- Hayes, R. A., Carr, C. T. & Wohn, D. Y. (2016). One Click, Many Meanings: Interpreting Paralinguistic Digital Affordances in Social Media. *Journal of Broadcasting & Electronic Media*, 60 (1), 171-187. Zugriff auf <http://dx.doi.org/10.1080/08838151.2015.1127248>
- Hecht, B. & Stephens, M. (2014). A Tale of Cities: Urban Biases in Volunteered Geographic Information. In *Proceedings of ICWSM'14*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8114>
- Heckner, M., Heilemann, M. & Wolff, C. (2009). Personal Information Management vs. Resource Sharing: Towards a Model of Information Behavior in Social Tagging Systems. In *Proceedings of ICWSM'09*. Washington: The AAAI Press. Zugriff auf <https://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/212>
- Hesterberg, T., Moore, D. S., Monaghan, S., Clipson, A., Epstein, R., Craig, B. A. & McCabe, B. A. (2010). Bootstrap Methods and Permutatios Tests. In G. P. M. David S. Moore & N. Bruce A. Craig W. H. Freeman (Hrsg.), *Introduction to the Practice of Statistics* (7. Aufl.). London: Freeman.
- Hider, P. & Pymm, B. (2008). Empirical research methods reported in high-profile {LIS} journal literature . *Library & Information Science Research*, 30 (2), 108 - 114. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0740818808000352>
- Hoang, T. A. (2016). *User Behavior Mining in Microblogging. Dissertations and Theses Collection (Open Access)*. Zugriff auf [http://ink.library.smu.edu.sg/etd\\_coll/130](http://ink.library.smu.edu.sg/etd_coll/130)
- Hodas, N. O. & Lerman, K. (2012). How Visibility and Divided Attention Constrain Social Contagion. In *Proceedings of the 2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust* (S. 249–257). Washington, DC, USA: IEEE Computer Society. Zugriff auf <http://dx.doi.org/10.1109/SocialCom-PASSAT.2012.129>
- Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, 30 (6). Zugriff auf <http://bst.sagepub.com/content/early/2010/10/31/0270467610385893>

- Holling, H. & Gediga, G. (2015). *Statistik - Testverfahren*. Göttingen: Hogrefe.
- Honey, C. & Herring, S. C. (2009). Beyond Microblogging: Conversation and Collaboration via Twitter. In *Proceedings of ICSS'09* (S. 1-10). Washington, DC, USA: IEEE. Zugriff auf <http://ieeexplore.ieee.org/document/4755499/>
- Hong, L., Dan, O. & Davison, B. D. (2011). Predicting Popular Messages in Twitter. In *Proceedings of WWW'11* (S. 57–58). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1963192.1963222>
- Hossain, N., Hu, T., Feizi, R., White, A. M., Luo, J. & Kautz, H. A. (2016). Inferring fine-grained details on user activities and home location from social media: Detecting drinking-while-tweeting patterns in communities. *CoRR*, *abs/1603.03181*. Zugriff auf <http://arxiv.org/abs/1603.03181>
- Howard, P. N. & Kollanyi, B. (2016). Bots, #StrongerIn, and #brexit: Computational propaganda during the UK-EU referendum. *CoRR*, *abs/1606.06356*. Zugriff auf <http://arxiv.org/abs/1606.06356>
- Howe, J. (2006). *Crowdsourcing: A Definition*. Zugriff auf [http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing\\_a.html](http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html)
- Howe, J. (2008). *Crowdsourcing. Why the power of the crowd is driving the future of business*. New York: Three River Press.
- Huang, J., Thornton, K. M. & Efthimiadis, E. N. (2010). Conversational Tagging in Twitter. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia* (S. 173–178). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1810617.1810647>
- Huang, L. (2016). *Moving Top Tweet Search Results from Reverse Chronological Order to Relevance Order*. Zugriff auf <https://blog.twitter.com/2016/moving-top-tweet-search-results-from-reverse-chronological-order-to-relevance-order>
- Hurlock, J. & Wilson, M. L. (2011). Searching Twitter: Separating the Tweet from the Chaff. In *Proceedings of ICWSM'11*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewFile/2819/3284>
- Höchstötter, N. (2007). Suchverhalten im Web - Erhebung, Analyse, Möglichkeiten. *Information Wissenschaft & Praxis*, 53 (3), 135–140.
- Ingwersen, P. & Järvelin, K. (2005). *The Turn*. Berlin: Springer.
- Ipeirotis, P. G., Provost, F. & Wang, J. (2010). Quality Management on Amazon Mechanical Turk. In *Proceedings of the ACM SIGKDD Workshop on Human Computation* (S. 64–67). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1837885.1837906>
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. Berlin: Springer.
- Jang, J. Y., Han, K. & Lee, D. (2015). No Reciprocity in "Liking"Photos: Analyzing Like Activities in Instagram. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media* (S. 273–282). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2700171.2791043>
- Jansen, B. J. (2006). Search log analysis: What it is, what's been done, how to do it. *Library & Information Science Research*, 28 (3), 407 - 432. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0740818806000673>
- Jansen, B. J. & Spink, A. (2006). How are we searching the World Wide Web? A comparison of nine search engine transaction logs. *Information Processing & Management*, 42 (1), 248 - 263. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0306457304001396>

- Jansen, B. J., Spink, A., Bateman, J. & Saracevic, T. (1998). Real Life Information Retrieval: A Study of User Queries on the Web. *SIGIR Forum*, 32 (1), 5–17. Zugriff auf <http://doi.acm.org/10.1145/281250.281253>
- Jansen, B. J., Spink, A., Blakely, C. & Koshman, S. (2007). Defining a Session on Web Search Engines: Research Articles. *Journal of the American Society for Information Science and Technology*, 58 (6), 862–871. Zugriff auf <http://dx.doi.org/10.1002/asi.v58:6>
- Jansen, B. J., Taksa, I. & Spink, A. (2009). Research and methodological foundations of transaction log analysis. In B. J. Jansen, I. Taksa & A. Spink (Hrsg.), *Handbook of Research on Web Log Analysis*. Hershey, PA: IGI Global.
- Java, A., Song, X., Finin, T. & Tseng, B. (2007). Why We Twitter: Understanding Microblogging Usage and Communities. In *Proceedings of WebKDD/SNA-KDD '07* (S. 56–65). ACM. Zugriff auf <http://doi.acm.org/10.1145/1348549.1348556>
- Jhaveri, N. & Räihä, K.-J. (2005). The Advantages of a Cross-session Web Workspace. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (S. 1949–1952). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1056808.1057064>
- Jiang, J., Wilson, C., Wang, X., Huang, P., Sha, W., Dai, Y. & Zhao, B. Y. (2010). Understanding Latent Interactions in Online Social Networks. In *Proceedings of sigcomm'10* (S. 369–382). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1879141.1879190>
- Jin, L., Chen, Y., Wang, T., Hui, P. & Vasilakos, A. V. (2013). Understanding user behavior in online social networks: a survey. *IEEE Communications Magazine*, 51 (9), 144–150.
- Joachims, T. (2002). Optimizing Search Engines Using Clickthrough Data. In *Proceedings of sigkdd'02* (S. 133–142). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/775047.775067>
- Johnson, R. B., Onwuegbuzie, A. J. & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, 1 (2), 112–133.
- Jones, W. (2004). Finders, keepers? The present and future perfect in support of personal information management. *First Monday*, 9 (3). Zugriff auf <http://firstmonday.org/ojs/index.php/fm/article/view/1123>
- Jones, W. (2007). How People Keep and Organize Personal Information. In W. Jones & J. Teevan (Hrsg.), *Personal Information Management* (S. 35–56). Washington: University of Washington Press.
- Jones, W. (2008). *Keeping Found Things Found*. San Francisco: Morgan Kaufmann.
- Jones, W. (2012). The Future of Personal Information Management, Part I: Our Information, Always and Forever. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 4 (1), 1–125. Zugriff auf <http://dx.doi.org/10.2200/S00411ED1V01Y201203ICR021>
- Jones, W. (2013). Transforming Technologies to Manage Our Information: The Future of Personal Information Management, Part II. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 5 (4), 1–179. Zugriff auf <http://dx.doi.org/10.2200/S00532ED1V01Y201308ICR028>
- Jones, W. (2015). Building a Better World with Our Information: The Future of Personal Information Management, Part 3. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 7 (4), 1–203. Zugriff auf <http://dx.doi.org/10.2200/S00653ED1V01Y201506ICR042>
- Jones, W. & Bruce, H. (2005). *A Report on the NSF-Sponsored Workshop on Personal Information Management*. Seattle, WA.



- Jones, W., Bruce, H. & Dumais, S. (2001). Keeping Found Things Found on the Web. In *Proceedings of CIKM'01* (S. 119–126). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/502585.502607>
- Jones, W., Munat, C. F., Bruce, H. & Foxley, A. (2005). The universal labeler: Plan the project and let your information follow. *Proceedings of the American Society for Information Science and Technology*, 42 (1). Zugriff auf <http://dx.doi.org/10.1002/meet.14504201185>
- Jones, W., Phuwanartnurak, A. J., Gill, R. & Bruce, H. (2005). Don't Take My Folders Away!: Organizing Personal Information to Get Things Done. In *Proceedings of chi'05 ea* (S. 1505–1508). ACM. Zugriff auf <http://doi.acm.org/10.1145/1056808.1056952>
- Jones, W. & Ross, B. H. (2007). Personal Information Management. In F. T. Durso (Hrsg.), *Handbook of applied Cognition* (S. 471–496). New York, NY, USA: John Wiley & Sons, Inc.
- Jones, W. & Teevan, J. (2007a). Introduction. In W. Jones & J. Teevan (Hrsg.), *Personal Information Management* (S. 3–20). Washington: University of Washington Press.
- Jones, W. & Teevan, J. (Hrsg.). (2007b). *Personal Information Management*. Washington: University of Washington Press.
- Jones, W., Wenning, A. & Bruce, H. (2014). How do people re-find files, emails and web pages? In *Proc. iConference'14* (S. 552–564). Zugriff auf <http://hdl.handle.net/2142/47300>
- Julien, H. (1996). A Content Analysis of the recent Information Needs and Uses Literature. *Library & Information Science Research*, 18 (1), 53 - 65. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0740818896900304>
- Julien, H. & Duggan, L. J. (2000). A Longitudinal Analysis of the Information Needs and Uses Literature. *Library & Information Science Research*, 22 (3), 291 - 309. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0740818899000572>
- Julien, H. & O'Brien, M. (2014). Information Behaviour Research: Where Have We Been, Where Are We Going? *Canadian Journal of Information and Library Science*, 38 (4), 239–250.
- Julien, H., Pecoskie, J. J. & Reed, K. (2011). Trends in information behavior research, 1999–2008: A content analysis. *Library & Information Science Research*, 33 (1), 19 - 24. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S074081881000112X>
- Järvelin, K. & Vakkari, P. (1993). The evolution of library and information science 1965–1985: A content analysis of journal articles. *Information Processing & Management*, 29 (1), 129 - 144. Zugriff auf <http://www.sciencedirect.com/science/article/pii/030645739390028C>
- Kabacoff, R. (2010). *R in Action*. Shelter Island, NY: Manning.
- Kallus, K. (2010). *Erstellung von Fragebogen*. Stuttgart: UTB.
- Kaplan, A. M. & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53 (1), 59 - 68. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0007681309001232>
- Kawase, R., Papadakis, G., Herder, E. & Nejd, W. (2011). Beyond the Usual Suspects: Context-aware Revisitation Support. In *Proceedings of the 22Nd ACM Conference on Hypertext and Hypermedia* (S. 27–36). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1995966.1995974>
- Khoo, C. (2014). Issues in information behaviour on social media. *LIBRES*, 24 (2), 75–96.
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P. & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54

- (3), 241 - 251. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0007681311000061>
- Kim, Y. S. & Yum, B.-J. (2011). Recommender system based on click stream data using association rule mining. *Expert Systems with Applications*, 38 (10), 13320 - 13327. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0957417411006816>
- Kirchhoff, S., Kuhnt, S., Lipp, P. & Schlawin, S. (2010). *Der Fragebogen* (5. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kivran-Swaine, F., Ting, J., Brubaker, J., Teodoro, R. & Naaman, M. (2014). Understanding Loneliness in Social Awareness Streams: Expressions and Responses. In *Proceedings of ICWSM'14*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8038>
- Kollanyi, B., Howard, P. N. & Woolley, S. C. (2016). Bots and Automation over Twitter during the U.S. Election. *Project Report on Computational Propaganda, 2016.4*. Zugriff auf [www.politicalbots.org](http://www.politicalbots.org)
- Komlodi, A., Soergel, D. & Marchionini, G. (2006). Search histories for user support in user interfaces. *JASIST*, 57 (6), 803–807. Zugriff auf <http://dx.doi.org/10.1002/asi.20297>
- Kooti, F., Aiello, L. M., Grbovic, M., Lerman, K. & Mantrach, A. (2015). Evolution of Conversations in the Age of Email Overload. In *Proceedings of the 24th International Conference on World Wide Web* (S. 603–613). Geneva, Switzerland: International World Wide Web Conferences Steering Committee. Zugriff auf <https://doi.org/10.1145/2736277.2741130>
- Körner, C., Benz, D., Hotho, A., Strohmaier, M. & Stumme, G. (2010). Stop Thinking, Start Tagging: Tag Semantics Emerge from Collaborative Verbosity. In *Proceedings of the 19th International Conference on World Wide Web* (S. 521–530). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1772690.1772744>
- Kotov, A., Bennett, P. N., White, R. W., Dumais, S. T. & Teevan, J. (2011). Modeling and analysis of cross-session search tasks. In *Proceedings of sigir'11* (S. 5–14). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2009916.2009922>
- Koufogiannakis, D., Slater, L. & Crumley, E. (2004). A Content Analysis of Librarianship Research. *Journal of Information Science*, 30 (3), 227-239. Zugriff auf <http://jis.sagepub.com/content/30/3/227.abstract>
- Krug, S. (2016). *Reactions Now Available Globally*. Zugriff auf <http://newsroom.fb.com/news/2016/02/reactions-now-available-globally/>
- Kuckartz, U. (2012). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*. Weinheim: Beltz.
- Kuckartz, U. (2014). *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*. Wiesbaden: Springer Fachmedien Wiesbaden. Zugriff auf [http://dx.doi.org/10.1007/978-3-531-93267-5\\_1](http://dx.doi.org/10.1007/978-3-531-93267-5_1)
- Kuckartz, U., Rädiker, S., Ebert, T. & Schehl, J. (2013). *Statistik - Eine verständliche Einführung* (2. Aufl.). Berlin: Springer.
- Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science*, 42 (5), 361.
- Kumar, A. (2015). *Hearts on Twitter*. Zugriff auf <https://blog.twitter.com/2015/hearts-on-twitter>
- Kurth, M. (1993). The limits and limitations of transaction log analysis. *Library Hi Tech*, 11 (2), 98-104. Zugriff auf <http://dx.doi.org/10.1108/eb047888>

- Kwak, H., Lee, C., Park, H. & Moon, S. (2010). What is Twitter, a Social Network or a News Media? In *Proceedings of WWW '10* (S. 591–600). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1772690.1772751>
- Kwasnik, B. (1989). How a Personal Document's Intended Use or Purpose Affects Its Classification in an Office. In *Proceedings of the 12th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 207–210). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/75334.75356>
- Kwasnik, B. (1991). The Importance of Factors that are not Document Attributes in the Organisation of Personal Documents. *Journal of Documentation*, 47 (4), 389–398. Zugriff auf <http://dx.doi.org/10.1108/eb026886>
- Lamnek, S. (2010). *Qualitative Sozialforschung*. Weinheim: Beltz.
- Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33 (1), 159–174.
- Lansdale, M. (1988). The psychology of personal information management. *Applied Ergonomics*, 19 (1), 55–66.
- Larsson, A. O. & Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14 (5), 729–747. Zugriff auf <http://dx.doi.org/10.1177/1461444811422894>
- Lazar, J., Feng, J. H. & Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*. New York, NY, USA: John Wiley & Sons, Inc.
- Leahey, C. (2014). *Throwback Thursday: The psychology behind its success*. Zugriff auf <http://fortune.com/2014/05/01/throwback-thursday-the-psychology-behind-its-success/>
- Lee, S.-Y., Hansen, S. S. & Lee, J. K. (2016). What Makes Us Click "Like" on Facebook? Examining Psychological, Technological, and Motivational Factors on Virtual Endorsement. *Comput. Commun.*, 73 (PB), 332–341. Zugriff auf <http://dx.doi.org/10.1016/j.comcom.2015.08.002>
- Leftheriotis, I. & Giannakos, M. N. (2014). Using social media for work: Losing your time or improving your work? *Computers in Human Behavior*, 31, 134 - 142. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S074756321300366X>
- Legewie, H. (2006). Review: Jörg Strübing (2004). Grounded Theory. Zur sozialtheoretischen und epistemologischen Fundierung des Verfahrens der empirisch begründeten Theoriebildung [Grounded Theory: Its Social-theoretical and Epistemological Foundation]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 7 (2). Zugriff auf <http://www.qualitative-research.net/index.php/fqs/article/view/105>
- Leonardi, P. M., Huysman, M. & Steinfield, C. (2013). Enterprise Social Media: Definition, History, and Prospects for the Study of Social Technologies in Organizations. *Journal of Computer-Mediated Communication*, 19 (1), 1–19. Zugriff auf <http://dx.doi.org/10.1111/jcc4.12029>
- Levordashka, A. & Utz, S. (2016). Ambient awareness: From random noise to digital closeness in online social networks. *Computers in Human Behavior*, 60, 147–154. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0747563216300899>
- Levordashka, A., Utz, S. & Ambros, R. (2016). What's in a Like? Motivations for Pressing the Like Button. In *Proceedings of ICWSM'16*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13022>

- Lewis, D. W. (1987). Research on the Use of Online Catalogs and Its Implications for Library Practice. *Journal of Academic Librarianship*, 13 (3), 152–157.
- Liang, S., Ren, Z., Weerkamp, W., Meij, E. & de Rijke, M. (2014). Time-Aware Rank Aggregation for Microblog Search. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management* (S. 989–998). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2661829.2661905>
- Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2 (3), 18–22. Zugriff auf <http://CRAN.R-project.org/doc/Rnews/>
- Lietz, H., Wagner, C., Bleier, A. & Strohmaier, M. (2014). When Politicians Talk: Assessing Online Conversational Practices of Political Parties on Twitter. In *Proceedings of ICWSM'14*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8069>
- Lim, B. H., Lu, D., Chen, T. & Kan, M.-Y. (2015). #Mytweet via Instagram: Exploring User Behaviour Across Multiple Social Networks. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (S. 113–120). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2808797.2808820>
- Lin, J. & Efron, M. (2013). Evaluation As a Service for Information Retrieval. *SIGIR Forum*, 47 (2), 8–14. Zugriff auf <http://doi.acm.org/10.1145/2568388.2568390>
- Lin, J. & Mishne, G. (2012). A Study of Churn in Tweets and Real-Time Search Queries. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4599>
- Lin, Y.-R., Margolin, D., Keegan, B., Baronchelli, A. & Lazer, D. (2013). #Bigbirds Never Die: Understanding Social Dynamics of Emergent Hashtags. In *Proceedings of ICWSM'13*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6083>
- Linder, R., Snodgrass, C. & Kerne, A. (2014). Everyday Ideation: All of My Ideas Are on Pinterest. In *Proceedings of CHI'14* (S. 2411–2420). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557273>
- Litt, E. & Hargittai, E. (2014). A bumpy ride on the information superhighway: Exploring turbulence online. *Computers in Human Behavior*, 36, 520 - 529. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0747563214002362>
- Litt, E., Spottswood, E., Birnholtz, J., Hancock, J. T., Smith, M. E. & Reynolds, L. (2014). Awkward Encounters of an „Other“ Kind: Collective Self-presentation and Face Threat on Facebook. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & #38; Social Computing* (S. 449–460). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2531602.2531646>
- Liu, G., Nguyen, T. T., Zhao, G., Zha, W., Yang, J., Cao, J., ... Chen, W. (2016). Repeat Buyer Prediction for E-Commerce. In *Proceedings of SIGKDD'16* (S. 155–164). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2939672.2939674>
- Liu, Z. & Jansen, B. J. (2012). Almighty Twitter, what are people asking for? *Proceedings of the American Society for Information Science and Technology*, 49 (1), 1–10. Zugriff auf <http://dx.doi.org/10.1002/meet.14504901134>
- Lozar-Manfreda, K., Bosnjak, M., Haas, I. & Vehovar, V. (2008). Web survey response rates compared to other modes - A meta-analysis. *International Journal of Market Research*, 50 (1), 79–104.

- Ma, Z., Sun, A. & Cong, G. (2012). Will This #Hashtag Be Popular Tomorrow? In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 1173–1174). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2348283.2348525>
- Ma, Z., Sun, A. & Cong, G. (2013). On predicting the popularity of newly emerging hashtags in Twitter. *Journal of the American Society for Information Science and Technology*, 64 (7), 1399–1410. Zugriff auf <http://dx.doi.org/10.1002/asi.22844>
- Macskassy, S. A. & Michelson, M. (2011). Why do People Retweet? Anti-Homophily Wins the Day! In *Proceedings of ICWSM'11*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2790>
- Malone, T. W. (1983). How Do People Organize Their Desks?: Implications for the Design of Office Information Systems. *ACM Trans. Inf. Syst.*, 1 (1), 99–112. Zugriff auf <http://doi.acm.org/10.1145/357423.357430>
- Mao, K., Capra, L., Harman, M. & Jia, Y. (2017). A survey of the use of crowdsourcing in software engineering. *Journal of Systems and Software*, 126, 57–84. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0164121216301832>
- Marchionini, G. (1995). *Information Seeking in Electronic Environments*. New York, NY, USA: Cambridge University Press.
- Markey, K. (2007). Twenty-five Years of End-user Searching, Part 1: Research Findings. *Journal of the American Society for Information Science and Technology*, 58 (8), 1071–1081. Zugriff auf <http://dx.doi.org/10.1002/asi.v58:8>
- Marshall, C. C. & Bly, S. (2005). Saving and Using Encountered Information: Implications for Electronic Periodicals. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 111–120). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1054972.1054989>
- Marshall, C. C. & Lindley, S. E. (2014). Searching for Myself: Motivations and Strategies for Self-search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 3675–3684). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557356>
- Marshall, C. C. & Shipman, F. M. (2014). An Argument for Archiving Facebook As a Heterogeneous Personal Store. In *Proceedings of JCDL'14* (S. 11–20). Piscataway, NJ, USA: IEEE Press. Zugriff auf <http://dl.acm.org/citation.cfm?id=2740769.2740772>
- Marshall, C. C. & Shipman, F. M. (2015). Exploring the Ownership and Persistent Value of Facebook Content. In *Proceedings of CSCW'15* (S. 712–723). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2675133.2675203>
- Martin, N., Lessmann, S. & Voß, S. (2008). Crowdsourcing: Systematisierung praktischer Ausprägungen und verwandter Konzepte. In M. Bichler et al. (Hrsg.), *Multikonferenz Wirtschaftsinformatik*. GITO-Verlag, Berlin. Zugriff auf <http://dblp.uni-trier.de/db/conf/mkwi/mkwi2008.html#MartinLV08>
- Marwick, A. & boyd, d. (2011). To See and Be Seen: Celebrity Practice on Twitter. *Convergence: The International Journal of Research into New Media Technologies*, 17 (2), 139–158.
- Massey, C., TenBrook, S., Tatum, C. & Whittaker, S. (2014). PIM and Personality: What Do Our Personal File Systems Say About Us? In *Proceedings of CHI'14* (S. 3695–3704). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557023>
- Mayring, P. (2010). *Qualitative Inhaltsanalyse* (11. Aufl.). Weinheim: Beltz.

- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Zugriff auf <http://www.ssoar.info/ssoar/handle/document/39517>
- McCreadie, R., Soboroff, I., Lin, J., Macdonald, C., Ounis, I. & McCullough, D. (2012). On Building a Reusable Twitter Corpus. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 1113–1114). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2348283.2348495>
- Meier, F. (2012). *Crowdsourced Information Architecture. Evaluation der Informationsarchitektur von Websites mithilfe von Crowdsourcing und asynchronen Remote-Usability-Tests*. Regensburg, Deutschland.
- Meier, F. (2015). Informationsverhalten in Social Media. *Information - Wissenschaft & Praxis*, 66 (1), 22–28.
- Meier, F. & Elswailer, D. (2014a). Personal Information Management and Social Networks Re-finding on Twitter. In *Proceedings of the 5th Information Interaction in Context Symposium* (S. 339–341). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2637002.2637058>
- Meier, F. & Elswailer, D. (2014b). Tweets I've Seen: Analysing Factors Influencing Re-finding Frustration on Twitter. In *Proceedings of the 5th Information Interaction in Context Symposium* (S. 287–290). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2637002.2637044>
- Meier, F. & Elswailer, D. (2016). Going Back in Time: An Investigation of Social Media Re-finding. In *Proceedings of SIGIR'16* (S. 355–364). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2911451.2911524>
- Meier, F., Elswailer, D. & Wilson, M. L. (2014). More than Liking and Bookmarking? Towards Understanding Twitter Favouriting Behaviour. In *Proceedings of ICWSM'14*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8094>
- Meister, J. D. & Sullivan, D. J. (1967). *Evaluation of user reactions to a prototype online information retrieval system* (Bericht). Canoga Park: NASA.
- Michael R. Chernick, R. A. L. (2008). *Bootstrap Methods. A Guide for Practitioners and Researchers*. New York, NY, USA: John Wiley & Sons, Inc.
- Mindruta, R. (2013). 5 Reasons People Favorite Tweets on Twitter. *Brandwatch Blog*. Zugriff auf <https://www.brandwatch.com/2013/10/5-reasons-why-twitter-users-favorite/>
- Mondal, M., Messias, J., Ghosh, S., Gummadi, K. P. & Kate, A. (2016). Forgetting in Social Media: Understanding and Controlling Longitudinal Exposure of Socially Shared Data. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)* (S. 287–299). Denver, CO: USENIX Association. Zugriff auf <https://www.usenix.org/conference/soups2016/technical-sessions/presentation/mondal>
- Morris, D., Ringel Morris, M. & Venolia, G. (2008). SearchBar: A Search-centric Web History for Task Resumption and Information Re-finding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 1207–1216). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1357054.1357242>
- Morris, M. R. & Horvitz, E. (2007). S3: Storable Shareable Search. In C. Baranauskas, P. Palanque, J. Abascal & S. D. J. Barbosa (Hrsg.), *Human-Computer Interaction – INTERACT 2007: 11th IFIP TC 13 International Conference, Rio de Janeiro, Brazil, September 10-14, 2007, Proceedings, Part*

- I (S. 120–123). Berlin: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-3-540-74796-3\\_13](http://dx.doi.org/10.1007/978-3-540-74796-3_13)
- Mukherjee, P. & Jansen, B. J. (2016). Second screen interaction analysis for IRL events: Phase-category investigation of the super bowl 2015 social soundtrack. In *7th International Conference on Information and Communication Systems (ICICS)* (S. 13-19). Washington, DC, USA: IEEE. Zugriff auf <http://ieeexplore.ieee.org/document/7476079/>
- Muralidharan, S., Rasmussen, L., Patterson, D. & Shin, J.-H. (2011). Hope for Haiti: An analysis of Facebook and Twitter usage during the earthquake relief efforts. *Public Relations Review*, 37 (2), 175 - 177. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0363811111000294>
- Murthy, D. & A.Longwell, S. (2013). Twitter and Disasters: The uses of Twitter during the 2010 Pakistan floods. *Information, Communication & Society*, 16 (6), 837-855. Zugriff auf <http://dx.doi.org/10.1080/1369118X.2012.696123>
- Naaman, M., Boase, J. & Lai, C.-H. (2010). Is It Really About Me? Message Content in Social Awareness Streams. In *Proceedings of CSCW '10* (S. 189–192). ACM. Zugriff auf <http://doi.acm.org/10.1145/1718918.1718953>
- Naveed, N., Gottron, T., Kunegis, J. & Alhadi, A. C. (2011). Searching Microblogs: Coping with Sparsity and Document Quality. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management* (S. 183–188). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2063576.2063607>
- Nichols, J., Mahmud, J. & Drews, C. (2012). Summarizing Sporting Events Using Twitter. In *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces* (S. 189–198). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2166966.2166999>
- Nour, M. M. (1985). A quantitative analysis of the research articles published in core library journals of 1980. *Library & Information Science Research*, 7 (1), 261 - 273.
- Obendorf, H., Weinreich, H., Herder, E. & Mayer, M. (2007). Web Page Revisitation Revisited: Implications of a Long-term Click-stream Study of Browser Usage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 597–606). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1240624.1240719>
- Odijk, D., White, R. W., Hassan Awadallah, A. & Dumais, S. T. (2015). Struggling and Success in Web Search. In *Proceedings of CIKM'15* (S. 1551–1560). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2806416.2806488>
- Oeldorf-Hirsch, A., Hecht, B., Morris, M. R., Teevan, J. & Gergle, D. (2014). To Search or to Ask: The Routing of Information Needs Between Traditional Search Engines and Social Networks. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (S. 16–27). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2531602.2531706>
- Oh, C., Lee, T., Kim, Y., Park, S. & Suh, B. (2016). Understanding Participatory Hashtag Practices on Instagram: A Case Study of Weekend Hashtag Project. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (S. 1280–1287). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2851581.2892369>
- Oh, K. E. (2013). The Process of Organizing Personal Information. *Advances In Classification Research Online*, 23 (1), 66–67.
- O'Neil, L. (2013). *The 5 Possible Things a Twitter Favorite Actually Signifies*. Zugriff auf <http://bulletmedia.com/article/what-does-a-twitter-favorite-actually-signify/>

- Otsuka, E., Wallace, S. A. & Chiu, D. (2014). Design and Evaluation of a Twitter Hashtag Recommendation System. In *Proceedings of the 18th International Database Engineering & Applications Symposium* (S. 330–333). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2628194.2628238>
- Ottoni, R., Casas, D. L., Pesce, J. P., Jr; W. M., Wilson, C., Mislove, A. & Almeida, V. (2014). Of Pins and Tweets: Investigating How Users Behave Across Image- and Text-Based Social Networks. In *Proceedings of ICWSM'14*. Washington: The AAAI Press. Zugriff auf <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8059/8139>
- Ounis, I., Macdonald, C., Lin, J. & Soboroff, I. (2011). Overview of the trec-2011 microblog track. In *Proceedings of TREC'11*. Zugriff auf <http://homepage.cs.uiowa.edu/~psriniva/courses/22C149/ppts/MICROBLOG.OVERVIEW.pdf>
- Owoputi, O., Connor, B., Dyer, C., Gimpel, K., Schneider, N. & Smith, N. A. (2013). Improved part-of-speech tagging for online conversational text with word clusters. In *Proceedings of NAACL-HLT'13* (S. 380–390). Zugriff auf <http://www.cs.cmu.edu/~ark/TweetNLP/owoputi+etal.naacl13.pdf>
- Palmer, J. C. & Stickland, J. (2016). *A beginner's guide to crowdsourcing*. Zugriff auf <http://www.apa.org/science/about/psa/2016/06/changing-minds.aspx>
- Pálsdóttir, A. (2014). Preferences in the use of social media for seeking and communicating health and lifestyle information. *Information Research*, 19 (4). Zugriff auf <http://InformationR.net/ir/19-4/paper642.html>
- Parker, E. B. & Paisley, W. J. (1966). Research for Psychologists at the Interface of the Scientist and his Information System. *American Psychologist*, 21 (11), 1061–1071.
- Parzen, M., Lipsitz, S., Ibrahim, J. & Klar, N. (2002). An Estimate of the Odds Ratio That Always Exists. *Journal of Computational and Graphical Statistics*, 11 (2), 420–436.
- Paul, S. A., Hong, L. & Chi, E. H. (2011). Is Twitter a Good address for Asking Questions? A Characterization Study. In *Proceedings of ICWSM'11*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2813>
- Paßmann, J. & Gerlitz, C. (2014). ‚Good‘ platform-political reasons for ‚bad‘ platform-data. Zur sozio-technischen Geschichte der Plattformaktivitäten Fav, Retweet und Like. *medialekontrolle.de*, 3.1. Zugriff auf <http://www.medialekontrolle.de/wp-content/uploads/2014/09/Passmann-Johannes-Gerlitz-Carolin-2014-03-01.pdf>
- Peer, E., Vosgerau, J. & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, 46 (4), 1023–1031. Zugriff auf <http://dx.doi.org/10.3758/s13428-013-0434-y>
- Peng, G., Nunes, J. & Annansingh, F. (2011). Investigating information systems with mixed-methods research. In *Proceedings of the iadis international workshop on information systems research trends, approaches and methodologies*. Zugriff auf <http://eprints.whiterose.ac.uk/74737/>
- Penniman, D. W. (2009). Historic Perspective of Log Analysis. In B. J. Jansen, I. Taksa & A. Spink (Hrsg.), *Handbook of Research on Web Log Analysis*. Hershey, PA: IGI Global.
- Pervin, N., Phan, T. Q., Datta, A., Takeda, H. & Toriumi, F. (2015). Hashtag Popularity on Twitter: Analyzing Co-occurrence of Multiple Hashtags. In G. Meiselwitz (Hrsg.), *Social Computing and Social Media: 7th International Conference, SCSM 2015, Held as Part of HCI International 2015, Los Angeles, CA, USA, August 2-7, 2015, Proceedings* (S. 169–182). Cham: Springer International Publishing. Zugriff auf [http://dx.doi.org/10.1007/978-3-319-20367-6\\_18](http://dx.doi.org/10.1007/978-3-319-20367-6_18)



- Peters, T. A. (1993). The history and development of transaction log analysis. *Library Hi Tech*, 11 (2), 41-66. Zugriff auf <http://dx.doi.org/10.1108/eb047884>
- Petrelli, D., Bowen, S. & Whittaker, S. (2014). Photo mementos: Designing digital media to represent ourselves at home. *International Journal of Human-Computer Studies*, 72 (3), 320 - 336. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S107158191300133X>
- Petrovic, S., Osborne, M. & Lavrenko, V. (2013). I wish I didn't say that! analyzing and predicting deleted messages in twitter. *CoRR*, abs/1305.3107. Zugriff auf <http://arxiv.org/abs/1305.3107>
- Petrovic, S., Osborne, M., McCreddie, R., Macdonald, C., Ounis, I. & Shrimpton, L. (2013). Can Twitter Replace Newswire for Breaking News? In *Proceedings of ICWSM'13*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6066>
- Porst, R. (2014). *Fragebogen* (4. Aufl.). Berlin: Springer.
- Posch, L., Wagner, C., Singer, P. & Strohmaier, M. (2013). Meaning As Collective Use: Predicting Semantic Hashtag Categories on Twitter. In *Proceedings of the 22Nd International Conference on World Wide Web* (S. 621-628). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2487788.2488008>
- Rao, J., Lin, J. & Efron, M. (2015). Reproducible Experiments on Lexical and Temporal Feedback for Tweet Search. In A. Hanbury, G. Kazai, A. Rauber & N. Fuhr (Hrsg.), *Advances in Information Retrieval: 37th European Conference on IR Research, ECIR 2015, Vienna, Austria, March 29 - April 2, 2015. Proceedings* (S. 755-767). Cham: Springer International Publishing. Zugriff auf [http://dx.doi.org/10.1007/978-3-319-16354-3\\_82](http://dx.doi.org/10.1007/978-3-319-16354-3_82)
- Ravasio, P., Schär, S. G. & Krueger, H. (2004). In Pursuit of Desktop Evolution: User Problems and Practices with Modern Desktop Systems. *ACM Trans. Comput.-Hum. Interact.*, 11 (2), 156-180. Zugriff auf <http://doi.acm.org/10.1145/1005361.1005363>
- Rice, R. E. & Borgman, C. L. (1983). The use of computer-monitored data in information science and communication research. *Journal of the American Society for Information Science*, 34 (4), 247-256. Zugriff auf <http://dx.doi.org/10.1002/asi.4630340404>
- Ringel Morris, M., Teevan, J. & Panovich, K. (2010). What Do People Ask Their Social Networks, and Why? A Survey Study of Status Message Q&A Behavior. In *Proceedings of CHI'10* (S. 1739-1748). ACM. Zugriff auf <http://doi.acm.org/10.1145/1753326.1753587>
- Rissland, E. L. & Ashley, K. D. (1986). Hypotheticals As Heuristic Device. In *Proceedings of the Workshop on Strategic Computing Natural Language* (S. 165-178). Stroudsburg, PA, USA: Association for Computational Linguistics. Zugriff auf <http://dx.doi.org/10.3115/1077146.1077167>
- Risso, V. G. (2016). Research methods used in library and information science during the 1970-2010. *New Library World*, 117 (1/2), 74-93. Zugriff auf <http://dx.doi.org/10.1108/NLW-08-2015-0055>
- Robertson, S. E. & Hancock-Beaulieu, M. M. (1992). On the Evaluation of IR Systems. *Information Processing and Management*, 28 (4), 457-466. Zugriff auf [http://dx.doi.org/10.1016/0306-4573\(92\)90004-J](http://dx.doi.org/10.1016/0306-4573(92)90004-J)
- Rogers, K. (2016). *Edit Tweets? Users Debate a Missing Twitter Feature*. Zugriff auf [http://www.nytimes.com/2016/01/07/technology/edit-tweets-users-debate-a-missing-twitter-feature.html?\\_r=0](http://www.nytimes.com/2016/01/07/technology/edit-tweets-users-debate-a-missing-twitter-feature.html?_r=0)

- Rogers, R. (2013). Debanalizing Twitter: The Transformation of an Object of Study. In *Proceedings of WebSci'13* (S. 356–365). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2464464.2464511>
- Romero, D. M., Meeder, B. & Kleinberg, J. (2011). Differences in the Mechanics of Information Diffusion Across Topics: Idioms, Political Hashtags, and Complex Contagion on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (S. 695–704). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1963405.1963503>
- Rosania, P. (2015). *While you were away...* Zugriff auf <https://blog.twitter.com/2015/while-you-were-away-0>
- Ross, J., Irani, L., Silberman, M. S., Zaldivar, A. & Tomlinson, B. (2010). Who Are the Crowdworkers?: Shifting Demographics in Mechanical Turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (S. 2863–2872). ACM. Zugriff auf <http://doi.acm.org/10.1145/1753846.1753873>
- Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior*, 43, 304 - 307. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S0747563214005809>
- Russell, D. M. & Lawrence, S. (2007). Search Everything. In W. Jones & J. Teevan (Hrsg.), *Personal Information Management* (S. 153–166). Washington: University of Washington Press.
- Russell-Rose, T., Lamantia, J. & Makri, S. (2014). Defining and Applying a Language for Discovery. In A. Nürnberger, S. Stober, B. Larsen & M. Detyniecki (Hrsg.), *Adaptive Multimedia Retrieval: Semantics, Context, and Adaptation: 10th International Workshop, AMR 2012, Copenhagen, Denmark, October 24-25, 2012, Revised Selected Papers* (S. 3–28). Cham: Springer International Publishing. Zugriff auf [http://dx.doi.org/10.1007/978-3-319-12093-5\\_1](http://dx.doi.org/10.1007/978-3-319-12093-5_1)
- Sadeghi, S., Blanco, R., Mika, P., Sanderson, M., Scholer, F. & Vallet, D. (2015). Predicting Re-finding Activity and Difficulty. In A. Hanbury, G. Kazai, A. Rauber & N. Fuhr (Hrsg.), *Proceedings of ECIR'15* (S. 715–727). Cham: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-3-319-16354-3\\_78](http://dx.doi.org/10.1007/978-3-319-16354-3_78)
- Sadeghi, S., Blanco, R., Mika, P., Sanderson, M., Scholer, F. & Vallet, D. (2017). Re-finding Behaviour in Vertical Domains. *Transaction on Information Systems*, 35 (3). (to appear)
- Sanderson, M. & Dumais, S. (2007). Examining Repetition in User Search Behavior. In *Proceedings of the 29th European Conference on IR Research* (S. 597–604). Berlin: Springer. Zugriff auf <http://dl.acm.org/citation.cfm?id=1763653.1763725>
- Saracevic, T. (2009). Information science. In M. J. Bates & M. N. Maack (Hrsg.), *Encyclopedia of Library and Information Science* (S. 2570–2586). New York: Taylor & Francis.
- Sarno, D. (2009a). *Jack Dorsey on the Twitter ecosystem, journalism and how to reduce reply spam. Part II.* Zugriff auf <http://latimesblogs.latimes.com/technology/2009/02/jack-dorsey-on.html>
- Sarno, D. (2009b). *Twitter creator Jack Dorsey illuminates the site's founding document. Part I.* Zugriff auf <http://latimesblogs.latimes.com/technology/2009/02/twitter-creator.html>
- Savolainen, R. (1995). Everyday life information seeking: Approaching information seeking in the context of “way of life”. *Library & Information Science Research*, 17 (3), 259 - 294. Zugriff auf <http://www.sciencedirect.com/science/article/pii/0740818895900489>

- Schneider, F., Feldmann, A., Krishnamurthy, B. & Willinger, W. (2009). Understanding Online Social Network Usage from a Network Perspective. In *Proceedings of IMC '09* (S. 35–48). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1644893.1644899>
- Schreier, M. (2012). *Qualitative Content Analysis in Practice*. Los Angeles: Sage.
- Schreier, M. (2014). Ways of Doing Qualitative Content Analysis: Disentangling Terms and Terminologies. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 15 (1). Zugriff auf <http://www.qualitative-research.net/index.php/fqs/article/view/2043>
- Schreier, M. & Echterhoff, G. (2013). *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*. Berlin: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-3-642-34362-9\\_10](http://dx.doi.org/10.1007/978-3-642-34362-9_10)
- Schryen, G., Wagner, G. & Benlian, A. (2015). Theory of Knowledge for Literature Reviews: An Epistemological Model, Taxonomy and Empirical Analysis of IS Literature. In *International Conference on Information Systems (ICIS)*.
- Schuth, A., Balog, K. & Kelly, L. (2015). Overview of the Living Labs for Information Retrieval Evaluation (LL4IR) CLEF Lab 2015. In J. Mothe et al. (Hrsg.), *Proceedings of CLEF'15* (S. 484–496). Cham: Springer International Publishing. Zugriff auf [http://dx.doi.org/10.1007/978-3-319-24027-5\\_47](http://dx.doi.org/10.1007/978-3-319-24027-5_47)
- Segal, R. B. & Kephart, J. O. (1999). MailCat: An Intelligent Assistant for Organizing e-Mail. In *Proceedings of the Third Annual Conference on Autonomous Agents* (S. 276–282). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/301136.301209>
- Selke, S. (2016). Einleitung. In S. Selke (Hrsg.), *Lifelogging: Digitale Selbstvermessung und Lebensprotokollierung zwischen disruptiver Technologie und kulturellem Wandel* (S. 1–21). Wiesbaden: Springer Fachmedien Wiesbaden. Zugriff auf [http://dx.doi.org/10.1007/978-3-658-10416-0\\_1](http://dx.doi.org/10.1007/978-3-658-10416-0_1)
- Severyn, A. & Moschitti, A. (2015). Learning to Rank Short Text Pairs with Convolutional Deep Neural Networks. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 373–382). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2766462.2767738>
- Shamma, D., Kennedy, L. & Churchill, E. (2010). Conversational Shadows: Describing Live Media Events Using Short Messages. In *Proceedings of ICWSM'10*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1506>
- Sharifi, B., Hutton, M.-A. & Kalita, J. K. (2010). Experiments in Microblog Summarization. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing* (S. 49–56). Washington, DC, USA: IEEE. Zugriff auf <http://dx.doi.org/10.1109/SocialCom.2010.17>
- Shen, X., Tan, B. & Zhai, C. (2007). Privacy Protection in Personalized Search. *SIGIR Forum*, 41 (1), 4–17. Zugriff auf <http://doi.acm.org/10.1145/1273221.1273222>
- Sherman, T. (2016). *Coming soon: express even more in 140 characters*. Zugriff auf <https://blog.twitter.com/express-even-more-in-140-characters>
- Shikano, S. (2010). Einführung in die Inferenz durch den nichtparametrischen Bootstrap. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 191–204). Wiesbaden: VS Verlag für Sozialwissenschaften. Zugriff auf [http://dx.doi.org/10.1007/978-3-531-92038-2\\_9](http://dx.doi.org/10.1007/978-3-531-92038-2_9)

- Shim, W. (2003). Using Handheld Computers in Information Seeking Research. *Journal of Education for Library and Information Science*, 44 (3/4), 258-265. Zugriff auf <http://www.jstor.org/stable/20764039>
- Shou, L., Wang, Z., Chen, K. & Chen, G. (2013). Sumblr: Continuous Summarization of Evolving Tweet Streams. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 533–542). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2484028.2484045>
- Shu, C. (2015). *Twitter Officially Launches Its “Retweet With Comment” Feature*. Zugriff auf <https://techcrunch.com/2015/04/06/retweetception/>
- Silverstein, C., Marais, H., Henzinger, M. & Moricz, M. (1999). Analysis of a Very Large Web Search Engine Query Log. *SIGIR Forum*, 33 (1), 6–12. Zugriff auf <http://doi.acm.org/10.1145/331403.331405>
- Simpson & Larson, R. R. (1991). Between Scylla and Charybdis: Subject Searching in the Online Catalog. *Advances in Librarianship*, 15, 175–236.
- Sin, S.-C. J. & Kim, K.-S. (2013). International students’ everyday life information seeking: The informational value of social networking sites. *Library & Information Science Research*, 35 (2), 107 - 116. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S074081881300008X>
- Skoric, M., Poor, N., Achananuparp, P., Lim, E. P. & Jiang, J. (2012). Tweets and Votes: A Study of the 2011 Singapore General Election. In *System Science (HICSS), 2012 45th Hawaii International Conference on* (S. 2583-2591).
- Sleeper, M., Balebako, R., Das, S., McConahy, A. L., Wiese, J. & Cranor, L. F. (2013). The Post That Wasn’t: Exploring Self-censorship on Facebook. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (S. 793–802). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2441776.2441865>
- Sleeper, M., Cranshaw, J., Kelley, P. G., Ur, B., Acquisti, A., Cranor, L. F. & Sadeh, N. (2013). „I Read My Twitter the Next Morning and Was Astonished “: A Conversational Perspective on Twitter Regrets. In *Proceedings of CHI’13* (S. 3277–3286). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2470654.2466448>
- Small, T. A. (2011). What the Hashtag? *Information, Communication & Society*, 14 (6), 872-895. Zugriff auf <http://dx.doi.org/10.1080/1369118X.2011.554572>
- Smith, S. M., Roster, C. A., Golden, L. L. & Albaum, G. S. (2016). A multi-group analysis of online survey respondent data quality: Comparing a regular USA consumer panel to MTurk samples. *Journal of Business Research*, 69 (8), 3139 - 3148. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S014829631500627X>
- Spencer, C. C. (1971). Random time sampling with self-observation for library cost studies: Unit costs of interlibrary loans and photocopies at a regional medical library. *Journal of the American Society for Information Science*, 22 (3), 153–160. Zugriff auf <http://dx.doi.org/10.1002/asi.4630220303>
- Spink, A., Jansen, B. J., Wolfram, D. & Saracevic, T. (2002). From E-Sex to E-Commerce: Web Search Changes. *IEEE Computer*, 35 (3), 107–109. Zugriff auf <http://dx.doi.org/10.1109/2.989940>
- Spink, A., Ozmutlu, S., Ozmutlu, H. C. & Jansen, B. J. (2002). U.S. versus European web searching trends. *SIGIR Forum*, 36 (2), 32–38. Zugriff auf <http://doi.acm.org/10.1145/792550.792555>

- Statista. (2014). *Prognose zur Anzahl der monatlich aktiven Twitter-Nutzer in Deutschland bis 2016*. Zugriff auf <https://de.statista.com/statistik/daten/studie/546761/umfrage/anzahl-der-monatlich-aktiven-twitter-nutzer-in-deutschland/>
- Stone, B. (2006). *Six More Twitter Updates*. Zugriff auf <https://blog.twitter.com/2006/six-more-twitter-updates>
- Stone, B. (2007a). *Are You Twittering @ Me?* Zugriff auf <https://blog.twitter.com/2007/are-you-twittering-me>
- Stone, B. (2007b). *Project Retweet: Phase One*. Zugriff auf <https://blog.twitter.com/2009/project-retweet-phase-one>
- Strauss, A. & Corbin, J. (1990). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Los Angeles: Sage.
- Strauss, A. & Corbin, J. (1996). *Grounded Theory: Grundlagen qualitativer Sozialforschung*. Weinheim: Beltz.
- Strübing, J. (2008). *Grounded Theory* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Strübing, J. (2013). *Qualitative Sozialforschung: Eine komprimierte Einführung für Studierende*. München: Oldenbourg Verlag.
- Strübing, J. (2014). *Grounded Theory* (3. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Su, Q. & Chen, L. (2015). A method for discovering clusters of e-commerce interest patterns using click-stream data. *Electronic Commerce Research and Applications*, 14 (1), 1 - 13. Zugriff auf <http://www.sciencedirect.com/science/article/pii/S1567422314000726>
- Sugimoto, C. R., Li, D., Russell, T. G., Finlay, S. C. & Ding, Y. (2011). The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using latent Dirichlet allocation. *Journal of the American Society for Information Science and Technology*, 62 (1), 185–204. Zugriff auf <http://dx.doi.org/10.1002/asi.21435>
- Suh, B., Hong, L., Piroli, P. & Chi, E. H. (2010). Want to Be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network. In *Proceedings of SOCIALCOM '10* (S. 177–184). Washington, DC, USA: IEEE. Zugriff auf <http://dx.doi.org/10.1109/SocialCom.2010.33>
- Sun, T., Zhang, M. & Mei, Q. (2013). Unexpected Relevance: An Empirical Study of Serendipity in Retweets. In *Proceedings of ICWSM'13*. Washington: The AAAI Press. Zugriff auf <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6129>
- Swan, L. & Taylor, A. S. (2005). Notes on Fridge Surfaces. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (S. 1813–1816). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1056808.1057029>
- Swan, M. (2012). Crowdsourced Health Research Studies: An Important Emerging Complement to Clinical Trials in the Public Health Research Ecosystem. *Journal of Medical Internet Research*, 14 (2), e46.
- Tan, C., Lee, L. & Pang, B. (2014). The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter. In *Proceedings of ACL'14* (S. 175–185). ACL. Zugriff auf <https://arxiv.org/pdf/1405.1438v1.pdf>
- Tao, K., Abel, F., Hauff, C. & Houben, G.-J. (2012a). Twinder: A Search Engine for Twitter Streams. In M. Brambilla, T. Tokuda & R. Tolksdorf (Hrsg.), *Web Engineering: 12th International Conference, ICWE 2012, Berlin, Germany, July 23-27, 2012. Proceedings* (S. 153–168). Berlin: Springer. Zugriff auf [http://dx.doi.org/10.1007/978-3-642-31753-8\\_11](http://dx.doi.org/10.1007/978-3-642-31753-8_11)

- Tao, K., Abel, F., Hauff, C. & Houben, G.-J. (2012b). What makes a tweet relevant for a topic. In *Making Sense of Microposts (# MSM2012) Whorkshop held at WWW'12* (S. 49–56). Zugriff auf [http://ceur-ws.org/Vol-838/paper\\_08.pdf](http://ceur-ws.org/Vol-838/paper_08.pdf)
- Tauscher, L. & Greenberg, S. (1997). Revisitation Patterns in World Wide Web Navigation. In *Proceedings of CHI'97* (S. 399–406). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/258549.258816>
- Teddlie, C. & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, 13 (1), 12–28.
- Teevan, J. (2007). The Re:Search Engine: Simultaneous Support for Finding and Re-finding. In *Proceedings of UIST'07* (S. 23–32). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1294211.1294217>
- Teevan, J., Adar, E., Jones, R. & Potts, M. (2006). History Repeats Itself: Repeat Queries in Yahoo's Logs. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 703–704). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1148170.1148326>
- Teevan, J., Adar, E., Jones, R. & Potts, M. A. S. (2007). Information Re-retrieval: Repeat Queries in Yahoo's Logs. In *Proceedings of SIGIR'07* (S. 151–158). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1277741.1277770>
- Teevan, J., Alvarado, C., Ackerman, M. S. & Karger, D. R. (2004). The Perfect Search Engine is Not Enough: A Study of Orienteering Behavior in Directed Search. In *Proceedings of CHI'04* (S. 415–422). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/985692.985745>
- Teevan, J., Capra, R. & Perez-Quinones, M. (2007). How People Find Personal Information. In W. Jones & J. Teevan (Hrsg.), *Personal Information Management* (S. 22–34). Washington: University of Washington Press.
- Teevan, J., Collins-Thompson, K., White, R. W., Dumais, S. T. & Kim, Y. (2013). Slow Search: Information Retrieval Without Time Constraints. In *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval* (S. 1:1–1:10). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2528394.2528395>
- Teevan, J., Ramage, D. & Morris, M. R. (2011). #TwitterSearch: A Comparison of Microblog Search and Web Search. In *Proceedings of WSDM'11* (S. 35–44). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1935826.1935842>
- Tesch, R. (1990). *Qualitative Research: Analysis Types and Software Tools*. London: The Falmer Press.
- Titcomb, J. (2016). *Twitter may finally let users edit their tweets*. Zugriff auf <http://www.telegraph.co.uk/technology/2016/12/30/twitter-may-finally-let-users-edit-tweets/>
- Truschkat, I., Kaiser, M. & Reinartz, V. (2005). A Recipe Book Approach to Research? Practical Suggestions for Using Grounded Theory in Dissertations and Thesis Projects. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 6 (2).
- Tumasjan, A., Sprenger, T., Sandner, P. & Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In *Proceedings of ICWSM'10*. Washington: The AAAI Press. Zugriff auf <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1441/1852>

- Tuomaala, O., Järvelin, K. & Vakkari, P. (2014). Evolution of library and information science, 1965–2005: Content analysis of journal articles. *Journal of the Association for Information Science and Technology*, 65 (7), 1446–1462. Zugriff auf <http://dx.doi.org/10.1002/asi.23034>
- Twitter. (2013). *What is a favorite?* Zugriff auf <http://web.archive.org/web/20131024065406/http://support.twitter.com/articles/14214-what-are-favorites>
- Twitter. (2016a). *Retweeting another Tweet.* Zugriff auf <https://support.twitter.com/articles/20169873>
- Twitter. (2016b). *Thank You! Love, Twitter.* Zugriff auf <https://blog.twitter.com/2016/thank-you-love-twitter>
- Twitter. (2016c). *Twitter milestones. A selection of memorable moments.* Zugriff auf <https://about.twitter.com/company/press/milestones>
- Twitter. (2016d). *Twitter Usage/Company Facts.* Zugriff auf <https://about.twitter.com/company>
- Twitter. (2016e). *Using hashtags on Twitter.* Zugriff auf <https://support.twitter.com/articles/49309>
- Twitter. (2016f). *Using Twitter search.* Zugriff auf <https://support.twitter.com/articles/132700>
- Tyler, S. K. & Teevan, J. (2010). Large Scale Query Log Analysis of Re-finding. In *Proceedings of WSDM '10* (S. 191–200). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1718487.1718512>
- Tyler, S. K., Wang, J. & Zhang, Y. (2010). Utilizing re-finding for personalized information retrieval. In *Proceedings of the 19th acm international conference on information and knowledge management* (S. 1469–1472). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1871437.1871649> doi: 10.1145/1871437.1871649
- Uysal, I. & Croft, W. B. (2011). User Oriented Tweet Ranking: A Filtering Approach to Microblogs. In *Proceedings of CIKM'11* (S. 2261–2264). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2063576.2063941>
- Wagner, K. (2015). *Hey Twitter Advertisers – Want More Clicks? Lose the Hashtag!* Zugriff auf <http://www.recode.net/2015/6/9/11563398/hey-twitter-advertisers-want-more-clicks-lose-the-hashtag>
- Wakeling, S., Clough, P. & Sen, B. (2014). Investigating the Potential Impact of Non-personalized Recommendations in the OPAC: Amazon vs. WorldCat.Org. In *Proceedings of the 5th Information Interaction in Context Symposium* (S. 96–105). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2637002.2637015>
- Walther, J. B., Van Der Heide, B., Hamel, L. M. & Shulman, H. C. (2009). Self-Generated Versus Other-Generated Statements and Impressions in Computer-Mediated Communication: A Test of Warranting Theory Using Facebook. *Communication Research*, 36 (2), 229–253. Zugriff auf <http://crx.sagepub.com/content/36/2/229.abstract>
- Wan-Chik, R., Clough, P. & Sanderson, M. (2013). Investigating religious information searching through analysis of a search engine log. *Journal of the American Society for Information Science and Technology*, 64 (12), 2492–2506. Zugriff auf <http://dx.doi.org/10.1002/asi.22945>

- Wang, G., Konolige, T., Wilson, C., Wang, X., Zheng, H. & Zhao, B. Y. (2013). You Are How You Click: Clickstream Analysis for Sybil Detection. In *Proceedings of SEC'13* (S. 241-256). USENIX Association. Zugriff auf <http://dl.acm.org/citation.cfm?id=2534766.2534788>
- Wang, G., Zhang, X., Tang, S., Zheng, H. & Zhao, B. Y. (2016). Unsupervised Clickstream Clustering for User Behavior Analysis. In *Proceedings of chi'16* (S. 225-236). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2858036.2858107>
- Wang, X., Bendersky, M., Metzler, D. & Najork, M. (2016). Learning to Rank with Selection Bias in Personal Search. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (S. 115-124). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2911451.2911537>
- Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G. & Cranor, L. F. (2011). „I Regretted the Minute I Pressed Share“: A Qualitative Study of Regrets on Facebook. In *Proceedings of the Seventh Symposium on Usable Privacy and Security* (S. 10:1-10:16). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2078827.2078841>
- Water, N. V. D., Surprenant, N., Genova, B. & Atherton, P. (1976). Research in information science: An assessment. *Information Processing & Management*, 12 (2), 117 - 123. Zugriff auf <http://www.sciencedirect.com/science/article/pii/030645737690042X>
- Wen, J. (2003). Post-Valued Recall web pages: User disorientation hits the big time. *IT & Society*, 1, 184-194.
- West, R., White, R. W. & Horvitz, E. (2013). From Cookies to Cooks: Insights on Dietary Patterns via Analysis of Web Usage Logs. In *Proceedings of www'13* (S. 1399-1410). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2488388.2488510>
- Whittaker, S., Bellotti, V. & Gwizdka, J. (2007). Everything through Email. In *Personal Information Management* (S. 167-190). Washington: University of Washington Press.
- Whittaker, S., Bergman, O. & Clough, P. (2010). Easy on that trigger dad: a study of long term family photo retrieval. *Personal and Ubiquitous Computing*, 14 (1), 31-43. Zugriff auf <http://dx.doi.org/10.1007/s00779-009-0218-7>
- Whittaker, S., Matthews, T., Cerruti, J., Badenes, H. & Tang, J. (2011). Am I Wasting My Time Organizing Email?: A Study of Email Refinding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 3449-3458). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1978942.1979457>
- Whittaker, S. & Sidner, C. (1996). Email Overload: Exploring Personal Information Management of Email. In *Proceedings of CHI'96* (S. 276-283). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/238386.238530>
- Williams, S. A., Terras, M. & Warwick, C. (2013a). How Twitter Is Studied in the Medical Professions: A Classification of Twitter Papers Indexed in PubMed. *Medicine 2.0* (2), e2.
- Williams, S. A., Terras, M. M. & Warwick, C. (2013b). What do people study when they study Twitter? Classifying Twitter related academic papers. *Journal of Documentation*, 69 (3), 384-410. Zugriff auf <http://dx.doi.org/10.1108/JD-03-2012-0027>
- Wilson, T. D. (1981). On User Studies and Information Needs. *Journal of Documentation*, 37 (1), 3-15. Zugriff auf <http://dx.doi.org/10.1108/eb026702>
- Wilson, T. D. (2000). Human information behavior. *Information Science*, 3 (2), 49-55. Zugriff auf <http://inform.nu/Articles/Vol3/v3n2p49-56.pdf>
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.



- Wollny, A. & Marx, G. (2009). Sozialforschung – Ausgangspunkte und Ansätze für eine forschende Allgemeinmedizin. *Zeitschrift für Allgemeinmedizin*.
- Wood, M. (2005). Bootstrapped Confidence Intervals as an Approach to Statistical Inference. *Organizational Research Methods*, 8 (4), 454-470.
- Woodruff, A. (2014). Necessary, Unpleasant, and Disempowering: Reputation Management in the Internet Age. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 149–158). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557126>
- Wu, S., Hofman, J. M., Mason, W. A. & Watts, D. J. (2011). Who Says What to Whom on Twitter. In *Proceedings of the 20th International Conference on World Wide Web* (S. 705–714). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1963405.1963504>
- Xu, J., Compton, R., Lu, T.-C. & Allen, D. (2014). Rolling Through Tumblr: Characterizing Behavioral Patterns of the Microblogging Platform. In *Proceedings of the 2014 ACM Conference on Web Science* (S. 13–22). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2615569.2615694>
- Xu, J.-M., Burchfiel, B., Zhu, X. & Bellmore, A. (2013). An Examination of Regret in Bullying Tweets. In *Proceedings of HLT-NAACL'13* (S. 697–702). Zugriff auf <http://pages.cs.wisc.edu/~jerryzhu/pub/regretnaacl13.pdf>
- Xu, W., Grishman, R., Meyers, A. & Ritter, A. (2013). A preliminary study of tweet summarization using information extraction. In *Proceedings of the workshop on language in social media (lasm 2013)* (S. 20–29). Zugriff auf <http://www.aclweb.org/anthology/W13-1103>
- Yang, J. & Counts, S. (2010). Predicting the Speed, Scale, and Range of Information Diffusion in Twitter. In *Proceedings of ICWSM'10*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1468/1896>
- Yang, J., Morris, M., Teevan, J., Adamic, L. & Ackerman, M. (2011). Culture Matters: A Survey Study of Social Q&A Behavior. In *Proceedings of ICWSM'11*. Washington: The AAAI Press. Zugriff auf <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2755>
- Yang, L., Sun, T., Zhang, M. & Mei, Q. (2012). We Know What @You #Tag: Does the Dual Role Affect Hashtag Adoption? In *Proceedings of the 21st International Conference on World Wide Web* (S. 261–270). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2187836.2187872>
- Yang, X., MacDonald, C. & Ounis, I. (2016). Using Word Embeddings in Twitter Election Classification. *CoRR*, abs/1606.07006. Zugriff auf <http://arxiv.org/abs/1606.07006>
- Yates, C. & Partridge, H. (2015). Citizens and social media in times of natural disaster: exploring information experience. *Information Research*, 20 (1). Zugriff auf <http://InformationR.net/ir/20-1/paper659.html>
- Ye, C. & Wilson, M. L. (2014). A User Defined Taxonomy of Factors That Divide Online Information Retrieval Sessions. In *Proceedings of the 5th Information Interaction in Context Symposium* (S. 48–57). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2637002.2637010>
- Zappavigna, M. (2015). Searchable talk: the linguistic functions of hashtags. *Social Semiotics*, 25 (3), 274-291. Zugriff auf <http://dx.doi.org/10.1080/10350330.2014.996948>
- Zarro, M. & Hall, C. (2012). Pinterest: Social Collecting for #Linking #Using #Sharing. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries* (S. 417–418). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2232817.2232919>

- Zarro, M., Hall, C. & Forte, A. (2013). Wedding dresses and wanted criminals: Pinterest. com as an infrastructure for repository building. In *Proceedings of ICWSM'13*. Washington: The AAAI Press.
- Zhang, J., Qu, Y., Cody, J. & Wu, Y. (2010). A Case Study of Micro-blogging in the Enterprise: Use, Value, and Related Issues. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 123–132). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/1753326.1753346>
- Zhang, Y., Jansen, B. J. & Spink, A. (2009). Time series analysis of a Web search engine transaction log. *Inf. Process. Manage.*, 45 (2), 230–245. Zugriff auf <http://dx.doi.org/10.1016/j.ipm.2008.07.003>
- Zhao, X., Lampe, C. & Ellison, N. B. (2016). The Social Media Ecology: User Perceptions, Strategies and Challenges. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (S. 89–100). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2858036.2858333>
- Zhao, X. & Lindley, S. E. (2014). Curation Through Use: Understanding the Personal Value of Social Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 2431–2440). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2556288.2557291>
- Zhao, X., Salehi, N., Naranjit, S., Alwaalan, S., Volda, S. & Cosley, D. (2013). The Many Faces of Facebook: Experiencing Social Media As Performance, Exhibition, and Personal Archive. In *Proceedings of CHI'13* (S. 1–10). New York, NY, USA: ACM. Zugriff auf <http://doi.acm.org/10.1145/2470654.2470656>
- Zhou, L., Wang, W. & Chen, K. (2016). Tweet Properly: Analyzing Deleted Tweets to Understand and Identify Regrettable Ones. In *Proceedings of the 25th International Conference on World Wide Web* (S. 603–612). Geneva, Switzerland: International World Wide Web Conferences Steering Committee. Zugriff auf <http://dx.doi.org/10.1145/2872427.2883052>
- Zimmer, M. & Proferes, N. J. (2014). A topology of Twitter research: disciplines, methods, and ethics. *Aslib Journal of Information Management*, 66 (3), 250–261. Zugriff auf <http://dx.doi.org/10.1108/AJIM-09-2013-0083>
- Zins, C. (2007). Conceptions of Information Science: Research Articles. *Journal of the American Society for Information Science and Technology*, 58 (3), 335–350. Zugriff auf <http://dx.doi.org/10.1002/asi.v58:3>



# Anhang



# A

## Twitter-Umfrage

### A.1. Fragebogen

Twitter Survey

1. Demographic Aspects

**1.1 I identify my gender as ...**

☐ Female   ☐ Male   ☐ I would rather not say

**1.2 When were you born?**

☐ After 1994  
☐ 1985-1994  
☐ 1975-1984  
☐ 1965-1974  
☐ Before 1965

**1.3 What is the highest level of education you achieved or working towards?**

☐ School Qualifications  
☐ Vocational Education  
☐ Bachelor's Degree  
☐ Master's Degree  
☐ Doctoral Degree

**1.4 How long has it been since you posted your *first* Tweet?**

☐ about a month ago  
☐ about half a year ago  
☐ about a year or more  
☐ about 2 years or more  
☐ about 5 years or more

**1.5 What other social networks do you use?**

☐ Facebook   ☐ Google+   ☐ LinkedIn   ☐ MySpace   ☐ Pinterest   ☐ Other(s)

**1.6 Which of the following are common means for you to access and use Twitter? (tick all that apply)**

☐ Twitter Website   ☐ Official Twitter Client for Smartphones   ☐ TweetDeck   ☐ Tweeterific   ☐ UberSocial   ☐ Twitter for Mac   ☐ Other(s)

**1.7 How often do you update your status message (post a tweet)?**

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Remember to log in to **Twitter** for your convenience :-)

**1.8 How many Tweets have you sent?**

Number of posted Tweets

**1.9 How many followers do you have?**

Number of followers

**1.10 How many people do you follow?**

Number of people you follow

Next Question Block

## 2. Favouriting Behaviour

10/37 questions answered

2.1 Are you aware that you can favourite tweets in Twitter?

☐ Yes ☐ No

Next Question Block

## 2. Favouriting Behaviour

11/37 questions answered

2.2 How many Tweets have you marked as favourite?

Number of favourited Tweets

You can use this [Link](#) to get your favourites counted !!!

2.3 How often do you mark a Tweet as favourite?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.4 Can you name reasons that lead you to favourite tweets?

2.5 Give an example of a tweet you recently favoured...

look in your Twitter favourites list and copy paste it...

2.6 How often do you examine your favourites list / favourites timeline?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.7 How often do you favourite *one of your own* Tweets?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.8 Can you name reasons that lead you to favourite *your own* tweets?

2.9 How often do you unfavourite a Tweet?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2.10 Can you name reasons that lead you to unfavourite Tweets?

Next Question Block

### 3. Preserving Tweets

20/37 questions answered

3.1 How often do you come across a Tweet, which you consider worth preserving for later?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Next Question Block

### 3. Preserving Tweets

21/37 questions answered

3.2 Can you name reasons why you consider Tweets to be worth preserving?

3.3 Which of the following best applies?

When you want to preserve a Tweet for later, do you...

	strongly disagree	disagree	neutral	agree	strongly agree
tend to mark it as favourite?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to retweet it, so you can look it up in your own posted Tweets when you need to?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
use some additional store (e.g. browser bookmark, software (Pocket, Evernote...), E-mail inbox, Word document, paper printout, RSS Feed)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to do nothing special with such Tweets?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3.4 How often do you preserve one of *your own* Tweets (a Tweet you posted yourself) for later?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Next Question Block

### 3. Preserving Tweets

24/37 questions answered

3.5 Which of the following best applies?



When you want to preserve one of *your own* Tweets (a Tweet you posted yourself) for later, do you...

	strongly disagree	disagree	neutral	agree	strongly agree
tend to mark it as favourite?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to retweet it, so you can look it up in your own posted tweets when you need to?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
use some additional store (e.g. browser bookmark, software (Pocket, Evernote...), E-mail inbox, Word document, paper printout, RSS Feed)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to do nothing special with such tweets?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Next Question Block

#### 4. Re-finding Tweets

25/37 questions answered

4.1 How often does it occur, that you want to find a previously seen Tweet again?

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Next Question Block

#### 4. Re-finding Tweets

26/37 questions answered

4.2 Give an example of a Tweet you recently tried to find again

look into Twitter for inspiration, find a Tweet and paste it here

4.3 Can you name reasons that motivate you to find previously seen Tweets again?

4.4 Which of the following best applies?

When you want to find a Tweet again, do you...

	strongly disagree	disagree	neutral	agree	strongly agree
tend to search for it in your own timeline?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to search for it in your favourites list?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to look in the persons timeline, you can remember tweeting it?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to use the Twitter searchbox and initiate a query?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to use a general purpose search engine (like Google)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to look in a particular store (e.g. browser bookmark, software	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

(Pocket, Evernote...), E-mail inbox, Word document, paper printout, RSS-Feed)					
---	--	--	--	--	--

**4.5 How often does it occur, that you want to find *one of your own* Tweets again?**

never	rarely	a few times a month	weekly	a few times a week	daily	multiple times per day
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**4.6 Which of the following best applies?**

When you want to find *one of your own* Tweets again, do you...

	strongly disagree	disagree	neutral	agree	strongly agree
tend to search for it in your own timeline?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to search for it in your favourites list?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to use the Twitter searchbox and initiate a query?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to use a general purpose search engine (like Google)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
tend to look in a particular store (e.g. browser bookmark, software (Pocket, Evernote...), E-mail inbox, Word document, paper printout, RSS-Feed)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**4.7 Can you name reasons that motivate you to find *your own* Tweets again?**

**4.8 How strong do you agree with these statements?**

When I want to find a tweet again...

	strongly disagree	disagree	neutral	agree	strongly agree
It is primarily because of the Tweets content (e.g. informative, entertaining etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is primarily because the Tweet contains a link to an external source	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is primarily because the Tweet contains an image or video	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
It is primarily because the Tweet was posted by a person I know personally	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Re-finding Tweets on Twitter is easy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**4.9 What would you assume is a typical time gap between reading a tweet for the first time and trying to refind it?**

within 1 hour	within a day	within a week	within a month	longer than a month
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**4.10 What would you assume is the longest time gap between reading a tweet and trying to refind it?**

35/37 questions answered

**5.1 How would you rate Twitter's current facilities/options for:**

	very poor	poor	ok	good	very good
Preserving interesting or useful information in Tweets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Finding previously seen Tweets again	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

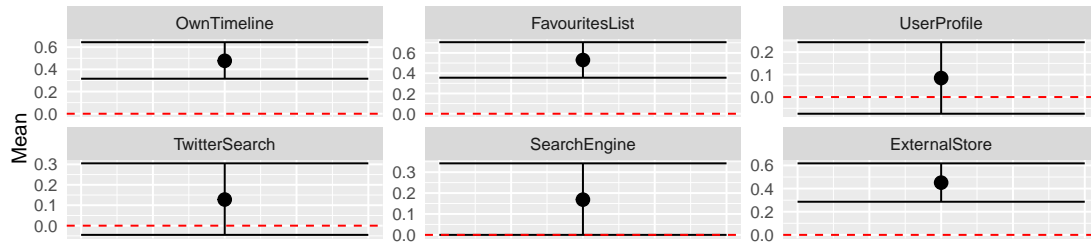
5.2 How often have you been frustrated, because you couldn't refind a previously seen Tweet?

[illegible]

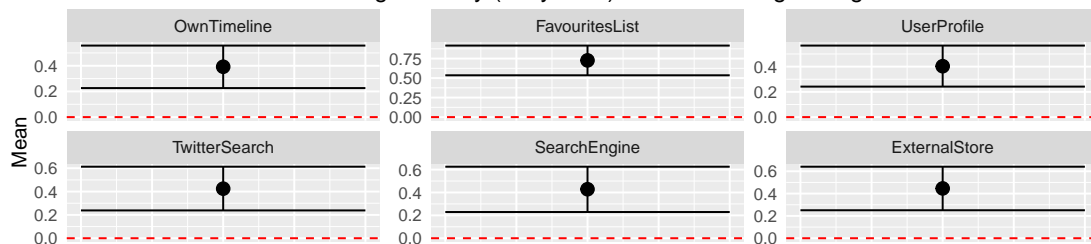
Finish Survey

## A.2. Signifikanzprüfungen der quantitativen Auswertungen des Fragebogens

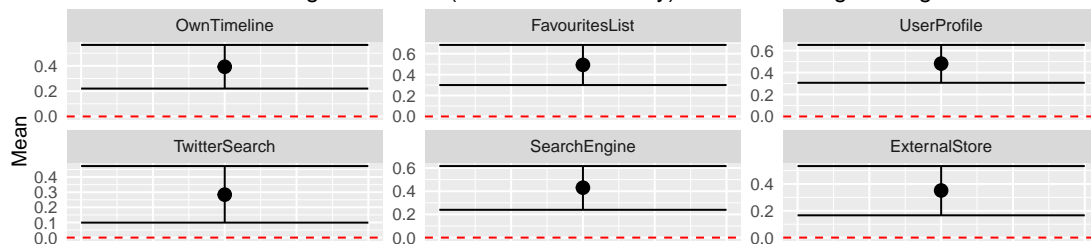
Preserving Frequency (never/at least rarely) and Re-finding Strategies



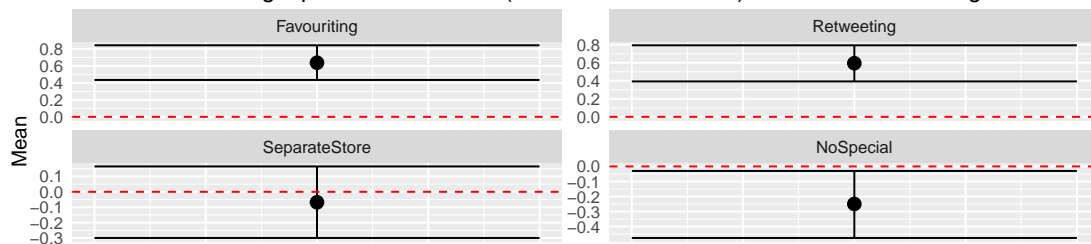
Re-finding Difficulty (easy/hard) and Re-finding Strategies



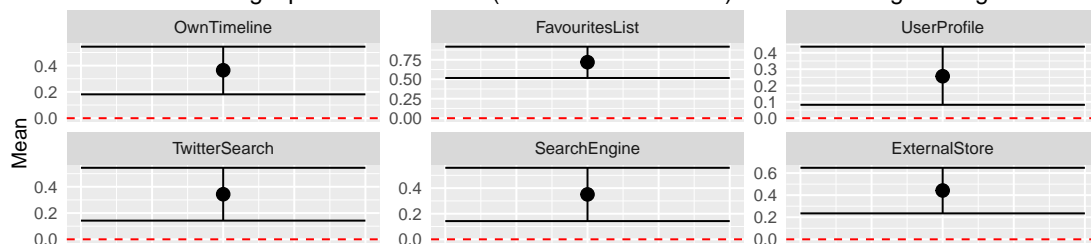
Re-finding Frustration (never/at least rarely) and Re-finding Strategies



Preserving Options Satisfaction (satisfied/not satisfied) and Preserve Strategies



Re-finding Options Satisfaction (satisfied/not satisfied) and Re-finding Strategies





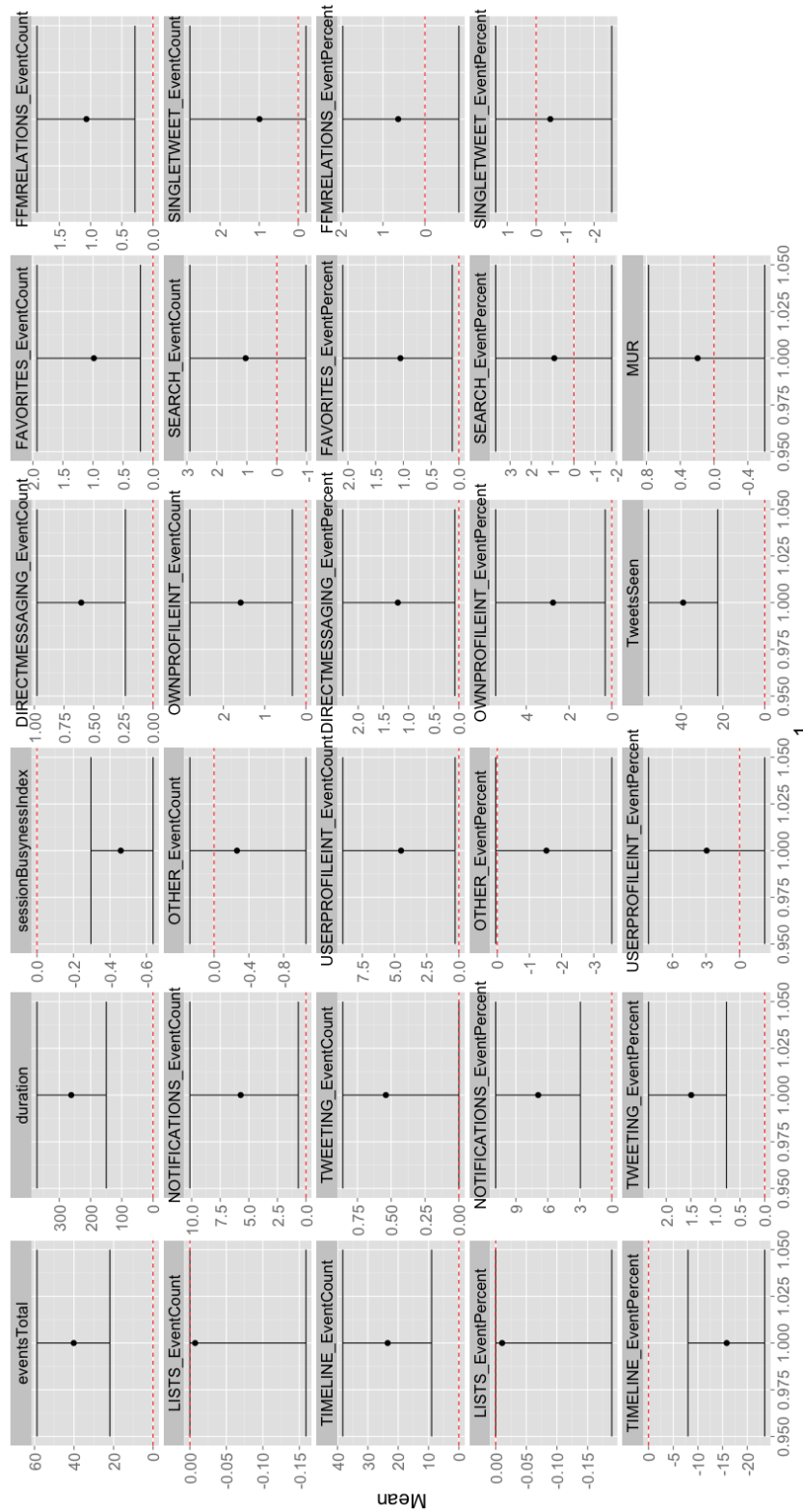
# Twitter-Logstudie

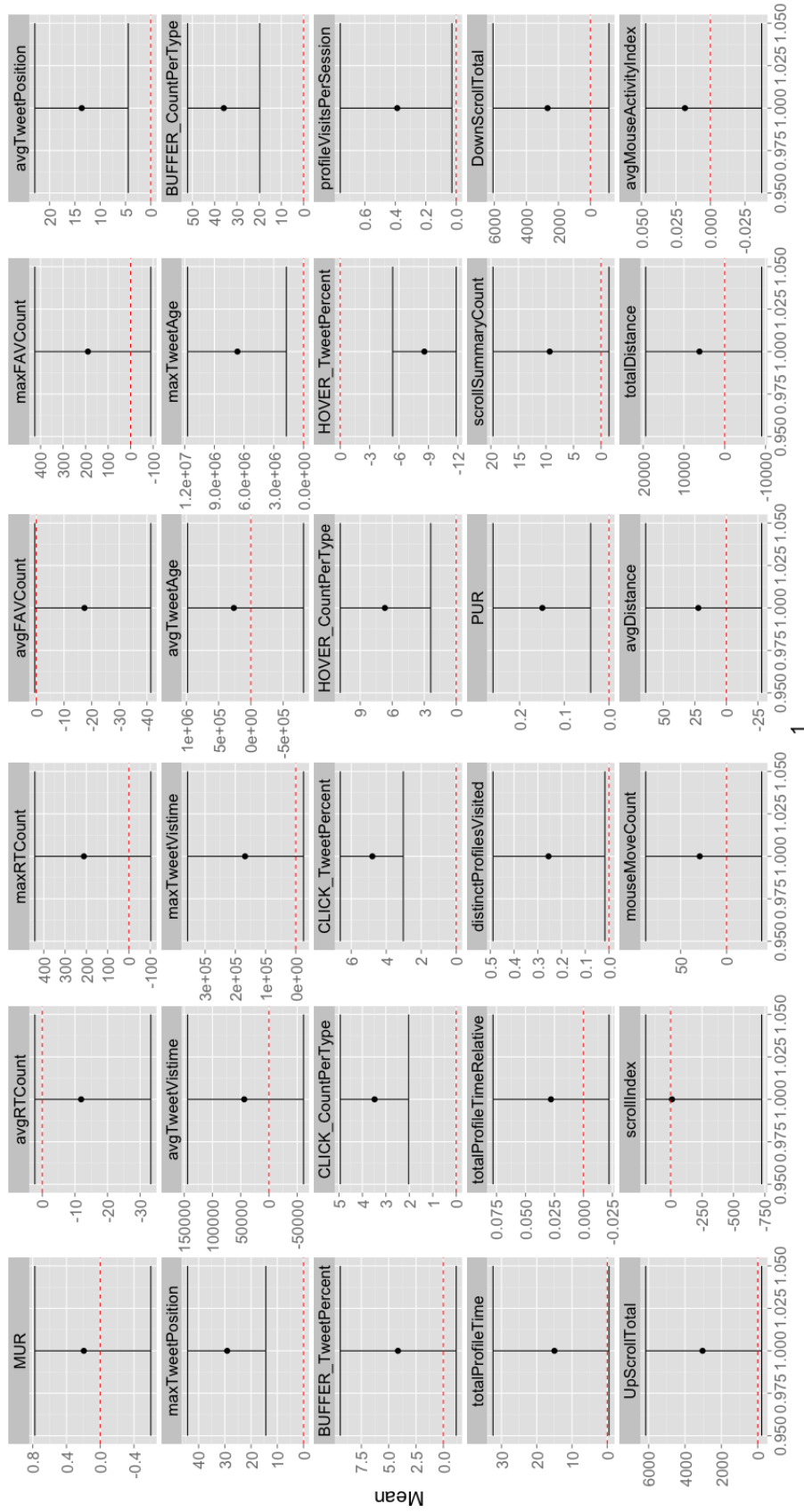
## B.1. Ergebnisse der Shapiro-Wilk-Tests

	W statistic	p.value
eventsTotal	0.45	0.00
duration	0.49	0.00
sessionBusynessIndex	0.61	0.00
DIRECTMESSAGING_EventCount	0.26	0.00
FAVORITES_EventCount	0.03	0.00
FFMRELATIONS_EventCount	0.18	0.00
LISTS_EventCount	0.00	0.00
NOTIFICATIONS_EventCount	0.11	0.00
OTHER_EventCount	0.08	0.00
OWNPROFILEINT_EventCount	0.08	0.00
SEARCH_EventCount	0.12	0.00
SINGLETWEET_EventCount	0.09	0.00
Timeline_EventCount	0.44	0.00
TWEETING_EventCount	0.10	0.00
USERPROFILEINT_EventCount	0.25	0.00
DIRECTMESSAGING_EventPercent	0.20	0.00
FAVORITES_EventPercent	0.03	0.00
FFMRELATIONS_EventPercent	0.26	0.00
LISTS_EventPercent	0.01	0.00
NOTIFICATIONS_EventPercent	0.36	0.00
OTHER_EventPercent	0.13	0.00
OWNPROFILEINT_EventPercent	0.14	0.00
SEARCH_EventPercent	0.21	0.00
SINGLETWEET_EventPercent	0.14	0.00
Timeline_EventPercent	0.68	0.00
TWEETING_EventPercent	0.18	0.00
USERPROFILEINT_EventPercent	0.41	0.00
TweetsSeen	0.57	0.00
MUR	0.19	0.00
avgRTCount	0.26	0.00
maxRTCount	0.22	0.00
avgFavCount	0.36	0.00
maxFavCount	0.27	0.00
avgTweetPosition	0.35	0.00
maxTweetPosition	0.45	0.00
avgTweetVistime	0.12	0.00
maxTweetVistime	0.23	0.00
avgTweetAge	0.11	0.00
maxTweetAge	0.33	0.00
BUFFERCountPerType	0.55	0.00
BUFFERTweetPercent	0.93	0.00
CLICKCountPerType	0.35	0.00
CLICKTweetPercent	0.49	0.00
HOVERCountPerType	0.47	0.00
HOVERTweetPercent	0.91	0.00
profileVisitsPerSession	0.28	0.00
totalProfileTime	0.26	0.00
totalProfileTimeRelative	0.36	0.00
distinctProfilesVisited	0.30	0.00
PUR	0.46	0.00
scrollSummaryCount	0.54	0.00
DownsScrollTotal	0.58	0.00
UpscrollTotal	0.38	0.00
scrollIndex	0.17	0.00
mouseMoveCount	0.62	0.00
avgDistance	0.55	0.00
totalDistance	0.48	0.00
avgMouseActivityIndex	0.29	0.00

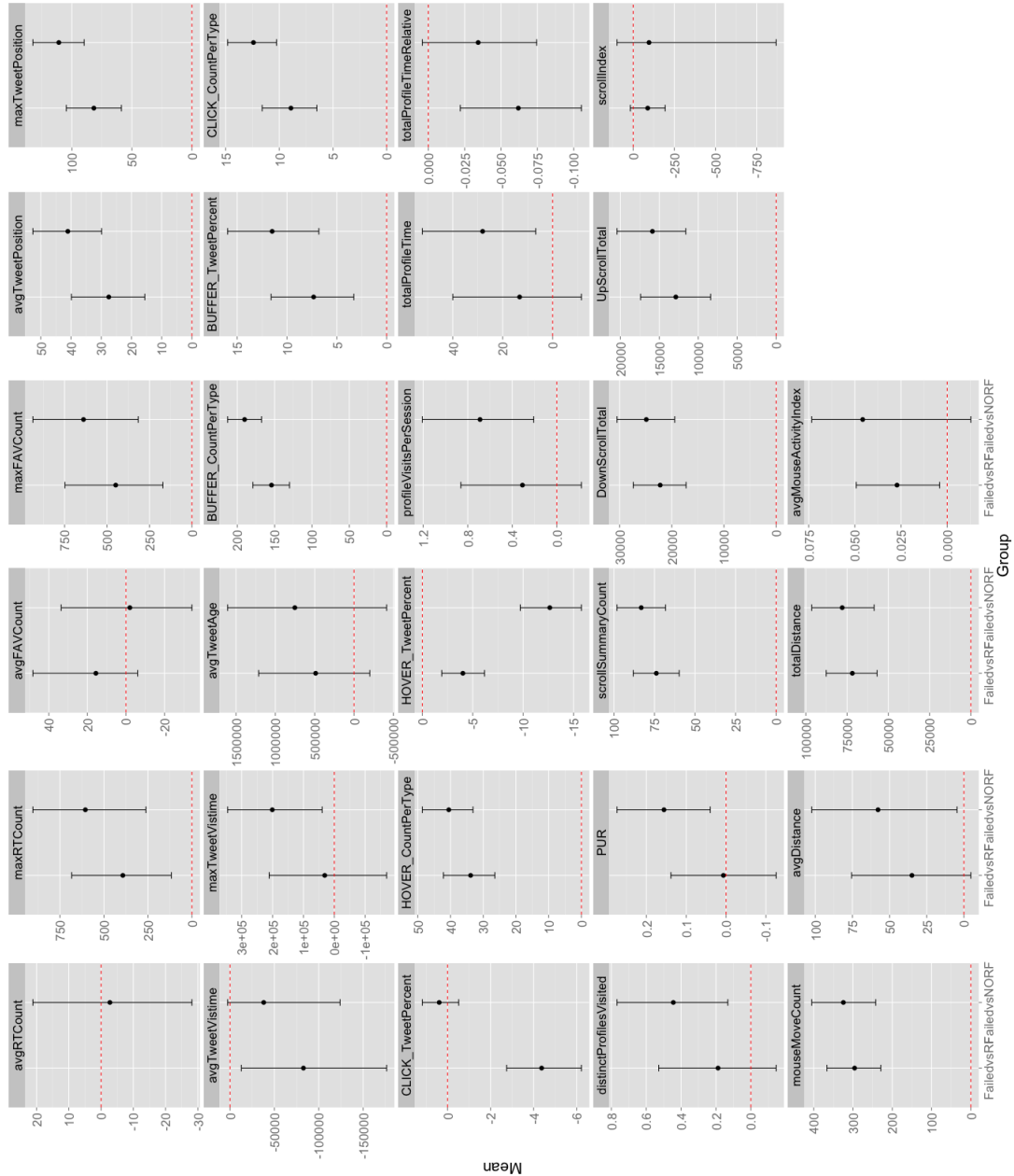
**Tabelle B.1.:** Verteilungstest der 58 Session-Variablen mit Hilfe von Shapiro-Wilk.

## B.2. Signifikanzprüfungen beim Vergleich von RF-Sessions mit normalen Sessions

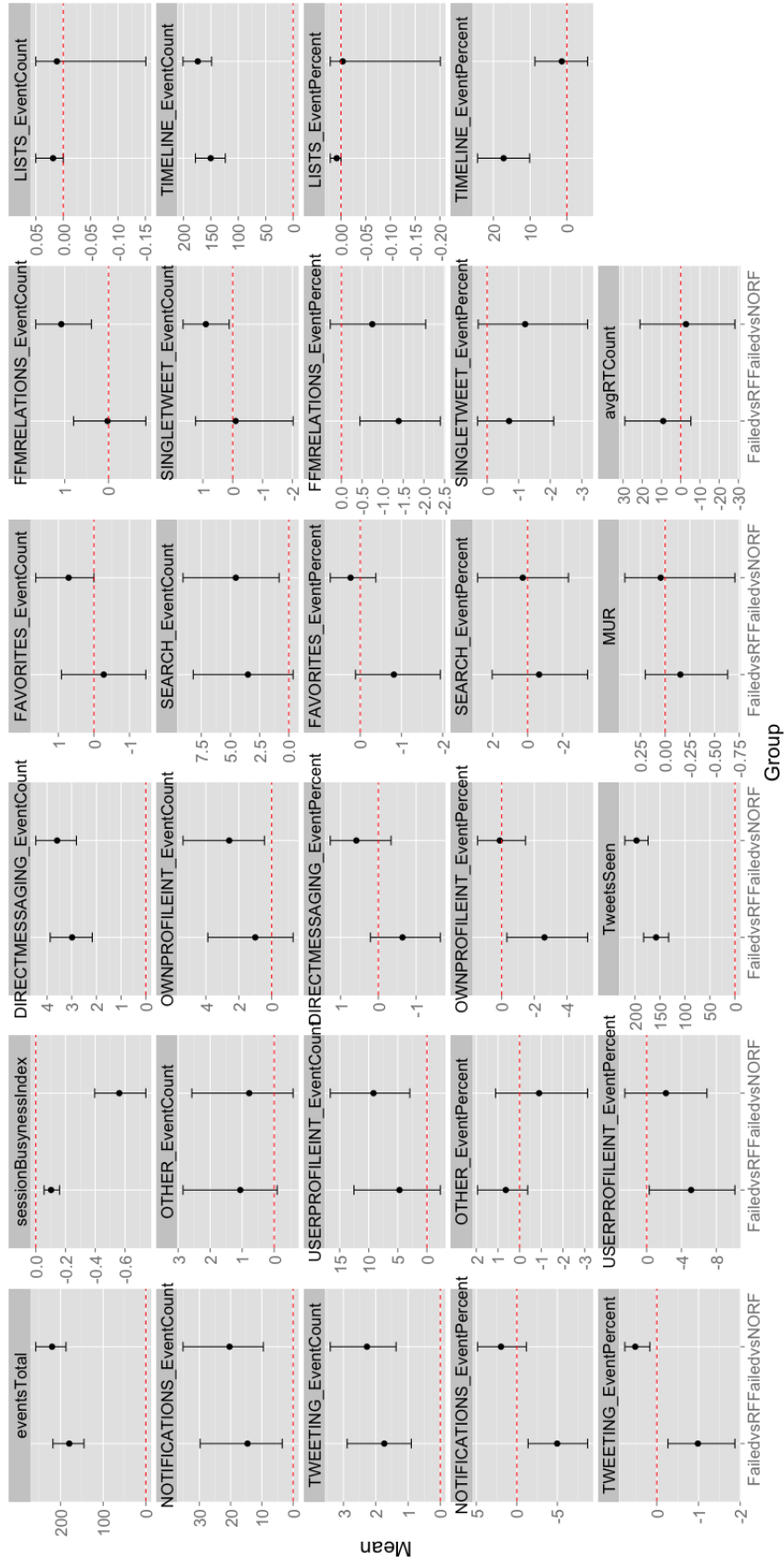




## B.3. Signifikanzprüfungen beim Vergleich von erfolgreichen und erfolglosen RF-Sessions mit normalen Sessions







## B.4. Kodierschema für Tweet-Inhaltskategorien

### Coding Schema for Tweet Classification

code	short name	explanation	example
CON	conversational	Tweets addressing one or multiple users. Questions directed to all followers, tweets referring to one or multiple persons.	<a href="#">@smashingmag</a> It isn't an ebook, but there's always the <a href="#">@A11YProject</a> <a href="#">a11yproject.com</a>
STA	status messages	Possible answers to Twitters old or new prompts („What are you doing“ „What is happening“ ) Me-now tweets referring to different situations in life (work, private) influenced by location and time, e.g. daily chatter	Joghurt mit Himbeeren ist Liebe und Liebe hat bekanntlich keine Kalorien.
NEW	news providing	Providing information, which is not user generated content – coverage of breaking news, discussion of live events e.g. weather forecast, sports results, natural disasters and so on.	Warum Forscher gerne gratis arbeiten - Open Source am Beispiel R <a href="#">science.orf.at/stories/176447...</a>
BRO	broadcasting and pass along	RTed tweets, content shared and/or generated by the user with links to websites/blog.	Awesome new <a href="#">#rstats</a> profiling tool from <a href="#">@winston_chang</a> ! Hadley Wickham added <b>Winston Chang</b> <a href="#">@winston_chang</a> Want to make your R code faster? Use <code>profvis</code> to find out what's making it slow. Then use your brain to speed it up. <a href="#">rpubs.com/wch/123888</a>
UNC	unclassifiable and spam	Unclassifiable strings of text or accidentally sent messages (cat-on-keyboard), junk and/or obviously automated posts, sponsored ads	:disapprove: :disapprove: :disapprove: :disapprove: :disapprove: :disapprove: :disapprove: :disapprove: :disapprove: :disapprove: